

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 15, 2004, 05:50:07 ; Search time 18600 Seconds
(without alignments)
11619.654 Million cell updates/sec

Title: US-09-921-143-36
Perfect score: 5283
Sequence: 1 aagcttgacctatgagct.....tgagtgtggcgcgcgtg 5283

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB	ID	Description
1	2410	45.6	6074	12	CVE18316	Y18316 Cloning vec
2	2240	42.4	8349	6	AR260588	AR260588 Sequence
3	2233.4	42.3	2661	12	SYNCKMRCG	M17626 Synthetic p
4	2218.4	42.0	2999	12	SYNCCDBA	L38498 Cloning vec
5	2218.4	42.0	2999	12	SYNCCDBB	L38499 Cloning vec
6	2177.4	41.2	2604	12	AY189826	AY189826 His-3 int
7	2177.4	41.2	11373	12	AY189827	AY189827 His-3 int
8	2177.4	41.2	11403	12	AY189829	AY189829 His-3 int
9	2088.4	39.5	3036	12	AY222815	AY222815 Cloning v
10	2088.4	39.5	3210	12	AY219859	AY219859 Cloning v
11	2088.4	39.5	5695	12	AY222822	AY222822 Shuttle v
12	2088.4	39.5	5869	12	AY219861	AY219861 Shuttle v
13	2088	39.5	5595	12	AF445080	AF445080 Cloning v
14	2078	39.3	3035	12	AY222814	AY222814 Cloning v
15	2078	39.3	3035	12	AY222811	AY222811 Cloning v
16	1892	35.8	7020	12	AS293724	Z93724 Murine retr
17	1793.4	33.9	4800	6	AR282049	AR282049 Sequence
18	1730.2	32.8	7414	12	AF402295	AF402295 PK[BIG-a]
19	1692.4	32.0	5225	6	AX226280	AX226280 Sequence
20	1671.2	31.6	5041	12	PKSM713	U04895 Cloning vec
21	1671.2	31.6	5041	12	PKSM715	U04896 Cloning vec
22	1671	31.6	4359	12	AY219686	AY219686 Expressio
23	1671	31.6	4968	12	AY219687	AY219687 Expressio
24	1669.6	31.6	7018	12	AY219683	AY219683 Shuttle e
25	1652.4	31.3	5594	6	ECO7659	AJ007659 Cloning v
26	1644.6	31.1	2959	12	AF402779	AF402779 Expressio
27	1644.6	31.1	2959	12	AF402779	AF402779 Expressio
28	1625.6	30.8	3741	12	PKSM711	U04894 Cloning vec
29	1618.2	30.6	6561	6	AR214683	AR214683 Sequence
30	1603.8	30.4	3599	6	BD063790	BD063790 Insulin-1
31	1603.8	30.4	3599	6	BD069041	BD069041 Treatment
32	1602.2	30.3	3600	6	BD063789	BD063789 Insulin-1
33	1602.2	30.3	3600	6	BD069040	BD069040 Treatment
34	1600	30.3	5966	6	BD136050	BD136050 Interfero
35	1572.6	29.8	8363	12	AF098284	AF098284 Cloning v
36	1571.6	29.7	4665	6	AR127825	AR127825 Sequence
37	1571.6	29.7	4665	6	BD082038	BD082038 Immunogen
38	1570.8	29.7	5063	6	BD139247	BD139247 Method of
39	1570	29.7	4517	6	E59418	E59418 Signal pept
40	1570	29.7	4518	6	AR214679	AR214679 Sequence
41	1570	29.7	4518	12	XXU37573	U37573 Shuttle exp
42	1570	29.7	4886	6	AR214349	AR214349 Sequence
43	1570	29.7	5619	6	AX008865	AX008865 Sequence
44	1570	29.7	5628	6	AX008867	AX008867 Sequence
45	1570	29.7	5775	6	AX641960	AX641960 Sequence

ALIGNMENTS

RESULT 1
CVE18316
LOCUS
DEFINITION Cloning vector pKGM.
ACCESSION Y18316
VERSION Y18316.1 GI:5830187
KEYWORDS bleomycin resistance; cloning vector; kanamycin resistance;
multiple cloning site; neomycin resistance.
SOURCE Cloning vector pKGM
ORGANISM Cloning vector pKGM
REFERENCE 1
AUTHORS Bannasch,D. and Schwab,M.
TITLE A versatile bait vector for rapid Gal4 dependent two-hybrid screens
JOURNAL Unpublished

3274 GATTTCACATCGCGTAGAATCCCGGAGGTGCTCCAGCTCAGGACGACTGAACAA 3333
3946 CTCGCGAGGGATCGAGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGG 4005
3334 CTCGCGAGGGATCGAGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGG 3393
4006 AGGATCATCCAGCCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAGAAGCG 4065
3394 AGGATCATCCAGCCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAGAAGCG 3453
4066 GCGGTGGAATCGAAATCTCGTGAAGCGAGGTGGCGCTCGCTTGGTGGTCAATTCGAAC 4125
3454 GCGGTGGAATCGAAATCTCGTGAAGCGAGGTGGCGCTCGCTTGGTGGTCAATTCGAAC 3513
4126 CCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAAGCGGATGCGCTGCGAAT 4185
3514 CCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAAGCGGATGCGCTGCGAAT 3573
4186 CGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTCAAGCCCATTCGCGCGCAAGCTCTT 4245
3574 CGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTCAAGCCCATTCGCGCGCAAGCTCTT 3633
4246 CAGCAATATCAGCGTAGCAACGCTATGCTGATAGCGGTCCGCCACACCCAGCGCGC 4305
3634 CAGCAATATCAGCGTAGCAACGCTATGCTGATAGCGGTCCGCCACACCCAGCGCGC 3693
4306 CACAGTGCATGAATCCAGAAAGCGCCATTTCCACCATCATATTCGGCAAGCAGCAT 4365
3694 CACAGTGCATGAATCCAGAAAGCGCCATTTCCACCATCATATTCGGCAAGCAGCAT 3753
4366 GCCATAGGTCAACGAGATCTCTCGCGTCGGGATCGCGGCTTGAGCGCTGCGGAACA 4425
3754 GCCATAGGTCAACGAGATCTCTCGCGTCGGGATCGCGGCTTGAGCGCTGCGGAACA 3813
4426 GTTCGCTGGCGAGCCCTGATGCTCTTGTCCAGATCATCTCTGATCGACACAGCGG 4485
3814 GTTCGCTGGCGAGCCCTGATGCTCTTGTCCAGATCATCTCTGATCGACACAGCGG 3873
4486 TTTCATCCGAGTACGTCGCTCGATGCGATGTTTCGCTGGTGGTTCGAATGGCGAGG 4545
3874 TTTCATCCGAGTACGTCGCTCGATGCGATGTTTCGCTGGTGGTTCGAATGGCGAGG 3933
4546 TAGCCGATCAAGCGTATGAGCGCCCGCATTCGATGATGATGATGATGATGATGATGAT 4605
3934 TAGCCGATCAAGCGTATGAGCGCCCGCATTCGATGATGATGATGATGATGATGATGAT 3993
4606 CAGGAGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4665
3994 CAGGAGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4053
4666 CCCTTCCCGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4725
4054 CCCTTCCCGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4113
4726 GCCAGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4785
4114 GCCAGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4737
4786 TGACAAAAGAACCGGGCGCCCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4845
4174 TGACAAAAGAACCGGGCGCCCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4233
4846 CGATTGCTGTTGTCAGCTCATAGCCGATAGCTCTTCCACCAAGCGCGCGCGGAAC 4905
4234 CGATTGCTGTTGTCAGCTCATAGCCGATAGCTCTTCCACCAAGCGCGCGCGGAAC 4293
4906 CTGCGTCAATCCATCTGTTCAATCATGCAAAACGATCTCTCTCTCTCTCTCTCTCTCTCT 4965
4294 CTGCGTCAATCCATCTGTTCAATCATGCAAAACGATCTCTCTCTCTCTCTCTCTCTCTCT 4353
4966 ATCTTGATCCCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5025
4354 ATCTTGATCCCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4413

QY 5026 AGGCTTTCCAACTTTACAGAGGGCGCCCGCTGCGCAATTCGGTTTCGCTTGTGCTCC 5085
Db 4414 AGGCTTTCCAACTTTACAGAGGGCGCCCGCTGCGCAATTCGGTTTCGCTTGTGCTCC 4473
QY 5086 ATAAACCCCGCTAGCTATGCGCATGTAAGCCCACTGCAAGCTACTGCTTCTCT 5145
Db 4474 ATAAACCCCGCTAGCTATGCGCATGTAAGCCCACTGCAAGCTACTGCTTCTCT 4533
QY 5146 TTGCGCTTTCGTTTTCCTTGTCCAGATAGCCAGTAGCTGACATTCATCCGGGGTCAGC 5205
Db 4534 TTGCGCTTTCGTTTTCCTTGTCCAGATAGCCAGTAGCTGACATTCATCCGGGGTCAGC 4593
QY 5206 ACCGTTTTCGCGACTGCGCTTCTAGCTTTCGCTTTCCTTTAGCAGCCCTTTCGCGCTG 5265
Db 4594 ACCGTTTTCGCGACTGCGCTTCTAGCTTTCGCTTTCCTTTAGCAGCCCTTTCGCGCTG 5265
QY 5266 AGTGTCTTTCGCGAGCGGTG 5283
Db 4654 AGTGTCTTTCGCGAGCGGTG 4671

RESULT 2
AR260588
LOCUS 8349 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 16 from patent US 6489542.
ACCESSION AR260588
VERSION AR260588.1 GI:27311143
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 8349)
AUTHORS Corbin,D.R. and Romano,C.P.
TITLE Methods for transforming plants to express Cry2ab
JOURNAL .delta.-endotoxins targeted to the plastids
Patent: US 6489542-A 16 03-DEC-2002;
Location/Qualifiers
source 1. .8349
BASE COUNT 1959 a 2355 c 2088 g 1947 t
ORIGIN

Query Match 42.4%; Score 2240; DB 6; Length 8349;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 25; Indels 144; Gaps 1;

QY 2706 CTAGAGCTTAATCATGCTCATAGCTTTCTCTGTTGAAATTCCTTATCCGCTCACAATTC 2765
Db 5911 CTGCGCTTAATCATGCTCATAGCTTTCTCTGTTGAAATTCCTTATCCGCTCACAATTC 5970
QY 2766 ACACAACTACAGCGGAGCATAAAGTAAAGCTGGGGTGCCTAATGAGTGAGCTA 2825
Db 5971 ACACAACTACAGCGGAGCATAAAGTAAAGCTGGGGTGCCTAATGAGTGAGCTA 6030
QY 2826 ACTCACTTAATGCTGCTGCTCATCTGCGCTTTCAGTCCGGAACCTGCTGCTGCA 2885
Db 6031 ACTCACTTAATGCTGCTGCTCATCTGCGCTTTCAGTCCGGAACCTGCTGCTGCA 6090
QY 2886 GCTGCACTTAATCAATCGCCCAACGCGGGAGAGCGGTTTTCGCTTATTCGGGCTCTTC 2945
Db 6091 GCTGCACTTAATCAATCGCCCAACGCGGGAGAGCGGTTTTCGCTTATTCGGGCTCTTC 6150
QY 2946 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
Db 6151 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6210
QY 3006 TCACTCAAGCGGTAATACGCTTATCCACAGAAACAGGGAATACGAGGAACAAAT 3065
Db 6211 TCACTCAAGCGGTAATACGCTTATCCACAGAAACAGGGAATACGAGGAACAAAT 6270
QY 3066 GTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAAAGCGCGCTTCTCGCTGCTTTT 3125


```
QY 4271 TATGTCCTGATAGCGGTCCGCCACACCCAGCGCGGCCACAGTCGATGAATCCAGAAAGCG 4330
Db 1027 TATGTCCTGATAGCGGTCCGCCACACCCAGCGCGGCCACAGTCGATGAATCCAGAAAGCG 368
QY 4331 GCCATTTTCCACCATGATATTCGCGACGAGGATCGCCATGGTTCAGGACGAGATCCCTC 4390
Db 967 GCCATTTTCCACCATGATATTCGCGACGAGGATCGCCATGGTTCAGGACGAGATCCCTC 308
QY 4391 GCCGTCGGGATGCGCGCTTGAGCTTGGCGAAACAGTTCGGCTGGCGGAGCCCTGATG 4450
Db 907 GCCGTCGGGATGCGCGCTTGAGCTTGGCGAAACAGTTCGGCTGGCGGAGCCCTGATG 848
QY 4451 CTCCTTCCTCAGATCCTGATCGACGACGAGCGCTTCCATCGGATGAGTCTCGCTC 4510
Db 847 CTCCTTCCTCAGATCCTGATCGACGAGCGCTTCCATCGGATGAGTCTCGCTC 788
QY 4511 GATCGGATGTTTCGCTTGGTGGTTCGATGAGCGGAGTAGCGGATCAAGCGTATGACGCG 4570
Db 787 GATCGGATGTTTCGCTTGGTGGTTCGATGAGCGGAGTAGCGGATCAAGCGTATGACGCG 728
QY 4571 CCGCATTTGCATCAGCCATGATGATCTTTCTCGGAGGAGCAAGTGATGATGACAGGAG 4630
Db 727 CCGCATTTGCATCAGCCATGATGATCTTTCTCGGAGGAGCAAGTGATGATGACAGGAG 668
QY 4631 ATCTGCGCCCGGCACTTCGCCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACACGTC 4690
Db 667 ATCTGCGCCCGGCACTTCGCCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACACGTC 608
QY 4691 GAGCAGAGTGGCGAAGAACGCGCTGCTGGCGAGCAGCAGATAGCGGCTGCTCGTC 4750
Db 607 GAGCAGAGTGGCGAAGAACGCGCTGCTGGCGAGCAGCAGATAGCGGCTGCTCGTC 548
QY 4751 CTGCAGTTTCATTCAGGCGACCGGACGCTCGTCTTGACAAAAGAACCGGCGCCCTG 4810
Db 547 TTGAGTTTCATTCAGGCGACCGGACGCTCGTCTTGACAAAAGAACCGGCGCCCTG 488
QY 4811 CGTGAAGCGGAAACGCGGATCAGAGCAGCGATGCTGTGTGTCGCGAGTCATA 4870
Db 487 CGTGAAGCGGAAACGCGGATCAGAGCAGCGATGCTGTGTGTCGCGAGTCATA 428
QY 4871 GCCCAATAGCTCTCCACCCAGCGGCGGAGAACCTGCTGCAATCCATCTGTTCAAT 4930
Db 427 GCCCAATAGCTCTCCACCCAGCGGCGGAGAACCTGCTGCAATCCATCTGTTCAAT 368
QY 4931 CATCGAAACGATCCTCATCTGCTCTTGATCAGATCTTGATCCCTCGGCGCATCAGAT 4990
Db 367 CATCGAAACGATCCTCATCTGCTCTTGATCAGATCTTGATCCCTCGGCGCATCAGAT 308
QY 4991 CTTTGGCGGCAAGAACGATCAGTTTACTTTTGCAGGGCTTCCCAACCTTACAGAGG 5050
Db 307 CTTTGGCGGCAAGAACGATCAGTTTACTTTTGCAGGGCTTCCCAACCTTACAGAGG 248
QY 5051 CGCCCGCAGCTGGCAATTCGGTTCGCTTGTGCTCCATAAAACCGCCAGTCTAGCTATCG 5110
Db 247 CGCCCGCAGCTGGCAATTCGGTTCGCTTGTGCTCCATAAAACCGCCAGTCTAGCTATCG 188
QY 5111 CCATGTAAGCCCATGCAAGCTACCTGCTTCTTTTGGCTTTCGCTTTTCCCTTCTCA 5170
Db 187 CCATGTAAGCCCATGCAAGCTACCTGCTTCTTTTGGCTTTCGCTTTTCCCTTCTCA 128
QY 5171 GATAGCCAGTAGCTGATTCATTCGCGGCTCAGCAGCTTCTTGGGAGTGGCTTCTA 5230
Db 127 GATAGCCAGTAGCTGATTCATTCGCGGCTCAGCAGCTTCTTGGGAGTGGCTTCTA 68
QY 5231 CGTGTTCGCTTCTTTAGCAGCCCTTGC9CCCTGAGTGTCTTGGCAGCGTG 5283
Db 67 CGTGTTCGCTTCTTTAGCAGCCCTTGC9CCCTGAGTGTCTTGGCAGCGTG 15
```

RESULT 4
SYNCCDBA/c 2999 bp DNA linear SYN 16-MAR-2000
LOCUS
DEFINITION Cloning vector pKil118 ccdB gene, complete cds and

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
source

kanamycin-resistance (KmR) gene fragment.
L38498
L38498.1 GI:986977
ccdB gene; cloning vector; kanamycin resistance.
Cloning vector pKil118
Cloning vector pKil118
artificial sequences; vectors.
1 (bases 1 to 2999)
Bernard,P.
New ccdB positive-selection cloning vectors with kanamycin or
chloramphenicol selectable markers
Gene 162 (1), 159-160 (1995)
96009896
7557407

Location/Qualifiers
1..2999
/organism="Cloning vector pKil118"
/mol_type="genomic DNA"
/db_xref="taxon:118286"
1..381
/gene="ccdB"
1..381
/gene="ccdB"
/codon_start=1
/transl_table=11
/protein_id="AAC42103.1"
/db_xref="GI:986977"
/translation="MTMITNSSSVPGDPLESTCRHASLSIQFKVYTKRSRVRFLFVD
VQSDIIDTPGRRMVPLASARLSDKVSRELYPVVHTGDESWRMTTDMASVPVSVIG
BEVADLSHRENDIKNAINLMFWGI"

gene

CDS

BASE COUNT 670 a 799 c 860 g 670 t
ORIGIN

Query Match 42.0%; Score 2218.4; DB 12; Length 2999;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 26; Indels 144; Gaps 1;

QY 2726 AGCTGTTTCTCTGTCGAAATGTTATCGCTCACAATTCACACAAATACGACGCCGAA 2785
Db 2999 AGCTGTTTCTCTGTCGAAATGTTATCGCTCACAATTCACACAAATACGACGCCGAA 2940
QY 2786 GCATAAGTGTAAGAGCTGGGTGCTTAATGAGTGAGCTAACTCACAATTAATTCGTTGC 2845
Db 2939 GCATAAGTGTAAGAGCTGGGTGCTTAATGAGTGAGCTAACTCACAATTAATTCGTTGC 2880
QY 2846 GCTCACTCCCGCTTTCAGTCGGGAAACCTGTCGCCAGCTGCTTAATGAATCGGCC 2905
Db 2879 GCTCACTCCCGCTTTCAGTCGGGAAACCTGTCGCCAGCTGCTTAATGAATCGGCC 2820
QY 2906 AACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCGGTTCCTCGTCACTGACT 2965
Db 2819 AACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCGGTTCCTCGTCACTGACT 2760
QY 2966 CGCTCGCTCGCTGCTCGGCTGGCGAGCGGATCAGCTCACTCAAGCGGTAATAC 3025
Db 2759 CGCTCGCTCGCTGCTCGGCTGGCGAGCGGATCAGCTCACTCAAGCGGTAATAC 2700
QY 3026 GGTATCCACAGAAATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGCGCAGAAA 3085
Db 2699 GGTATCCACAGAAATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGCGCAGAAA 2640
QY 3086 AGGCAGGAAACCGTAAGGCGCGGTTGCTGGGCTTTTCCATAGAGTCCGCCCTG 3145
Db 2639 AGGCAGGAAACCGTAAGGCGCGGTTGCTGGGCTTTTCCATAGAGTCCGCCCTG 2580
QY 3146 ACGAGCATCAAAAAATCGAGCTCAAGTCAGAGTGCGGCGGAAACCCGACAGGACTATAA 3205
Db 2579 ACGAGCATCAAAAAATCGAGCTCAAGTCAGAGTGCGGCGGAAACCCGACAGGACTATAA 2520
QY 3206 GATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGC 3265
Db 2519 GATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGC 2460


```
ORGANISM      Cloning vector pKIL119
REFERENCE      artificial sequences; vectors.
AUTHORS        1 (bases 1 to 2999)
TITLE          Bernard, P.
JOURNAL        New ccdB positive-selection cloning vectors with kanamycin or
MEDLINE        chloramphenicol selectable markers
PUBMED         Gene 162 (1), 159-160 (1995)
FEATURES       96009896
               7557407
               Location/Qualifiers
               1..2999
               /organism="Cloning vector pKIL119"
               /mol_type="genomic DNA"
               /db_xref="taxon:118287"
               1..381
               /gene="ccdB"
               1..381
               /gene="ccdB"
               /codon_start=1
               /transl_table=11
               /protein_id="AAC42104.1"
               /db_xref="GI:986980"
               /translation="MTMITPSLHACRSTLEDPVPSSNSLQFKVYTYKRSRVLFDV
               VQSDIIDPGRWVTPLASARILSKVSRELYPVVHIGDESWRMTTDMASVPVSVIG
               BEVADLSHRENDIKNAINLFWGI"
BASE COUNT    672 a 801 c 857 g 669 t
ORIGIN
Query Match      42.0%; Score 2218.4; DB 12; Length 2999;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 26; Indels 144; Gaps 1;
QY 2726 AGCTGTTTCCTGTGGAATTTGTTATCCGTCACAAATCCACAAATACGACCGGAA 2795
DB |-----|
QY 2786 GCATAAAGTGAAGCCTGGGTGCTTAATGAGTAGCTAACTACATTAATTCGTTGC 2845
DB |-----|
QY 2939 GCATAAAGTGAAGCCTGGGTGCTTAATGAGTAGCTAACTACATTAATTCGTTGC 2880
QY 2846 GCTCACTGCCGCTTCCAGTCGGGAACCTGTGTCGCGAGCTGATTAATGAATCGCC 2905
DB |-----|
QY 2879 GCTCACTGCCGCTTCCAGTCGGGAACCTGTGTCGCGAGCTGATTAATGAATCGCC 2920
QY 2906 AACGCGCGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCCGCTTCCCTCGCTCACTGACT 2965
DB |-----|
QY 2819 AACGCGCGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCCGCTTCCCTCGCTCACTGACT 2760
QY 2966 CGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGTCACTCAAGAGCGGTATAC 3025
DB |-----|
QY 2759 CGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGTCACTCAAGAGCGGTATAC 2700
QY 3026 GGTATCCACAGAAATCAGGGGATACGAGGAAAGAAATGAGCAAAAGGCCAGCAAA 3085
DB |-----|
QY 2699 GGTATCCACAGAAATCAGGGGATACGAGGAAAGAAATGAGCAAAAGGCCAGCAAA 2640
QY 3086 AGCCAGGAACCGTAAAAAGCGCGTTCGCTGGCGTTTTCATAGGCTTCGCGCCCGCTG 3145
DB |-----|
QY 2639 AGCCAGGAACCGTAAAAAGCGCGTTCGCTGGCGTTTTCATAGGCTTCGCGCCCGCTG 2580
QY 3146 ACAGCATCACAAAATCCAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAA 3205
DB |-----|
QY 2579 ACAGCATCACAAAATCCAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAA 2520
QY 3206 GATACAGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTTCCTGTTCCGACCTGCGCG 3265
DB |-----|
QY 2519 GATACAGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTTCCTGTTCCGACCTGCGCG 2460
QY 3266 TTACCGGATACCTGTCGCGCTTTCCTTCGCGAGCGTGGCGTTTCTCACTAGCTCAC 3325
DB |-----|
QY 2459 TTACCGGATACCTGTCGCGCTTTCCTTCGCGAGCGTGGCGTTTCTCACTAGCTCAC 2400
QY 3326 GCTGTAGTATCTCAGTTCGGTGTAGTGTGTTCCGCTCCAGGCTGGGCTGTGTGCAAGAAC 3385
DB |-----|
2399 GCTGTAGTATCTCAGTTCGGTGTAGTGTGTTCCGCTCCAGCTGGGCTGTGTGCAAGAAC 2340
3386 CCCCCGTTCCAGCCCGACCGCTCGGCTTATCCGTAACCTATCGTCTTGAGTCCAAACCCGG 3445
2339 CCCCCGTTCCAGCCCGACCGCTCGGCTTATCCGTAACCTATCGTCTTGAGTCCAAACCCGG 2280
3446 TAAGACACGACTTATCGCCACTCGGACGAGCAGCTGTTGAAGAGTACAGGATTAGCAGAGGAGT 3505
2279 TAAGACACGACTTATCGCCACTCGGACGAGCAGCTGTTGAAGAGTACAGGATTAGCAGAGGAGT 2220
3506 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGAA 3565
2219 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGAA 2160
3566 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAGAGTGTGTAGCT 3625
2159 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAGAGTGTGTAGCT 2100
3626 CTTGATCCGCAAAACAAACACCGCTGTAGCGGTGTTTTCGTTTTCGAAAGCAGCAGA 3685
2099 CTTGATCCGCAAAACAAACACCGCTGTAGCGGTGTTTTCGTTTTCGAAAGCAGCAGA 2040
3686 TTACGCGCAGAAAAAAGGATCTCAAGAAAGTCTTTTGAATCTTTTCTACGCGGTCTGAGC 3745
2039 TTACGCGCAGAAAAAAGGATCTCAAGAAAGTCTTTTGAATCTTTTCTACGCGGTCTGAGC 1980
3746 CTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGTCTATGAGATTATCTCGACCAAG 3805
1979 CTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGTCTATGAGATTATCAAAAGGATCT 1920
3806 CGGCCATCTGCTGCTCCCACTCTCCAGTTCGGGGGCGATGCGCGGATAGCCGCTGC 3865
1919 TCACCTAGATCCTT----- 1906
3866 TGGTTTCCTGGATGCGCAGCGGATTTGCATGCGGTAGAACTCCGCGAGGTGCTCCAGCC 3925
1905 ----- 1906
3926 TCAGGCGAGCGCTGAACCAACTCGCGAGGGATCGAGCCCGGGTGGCGAAGAACTCCA 3985
1905 -----TTGGGTGGCGAAGAACTCCA 1884
3986 GCATGAGATCCCGCGCTCGAGGATCATCCAGCCGCGTCCCGGAAAAAGATTCGGAAGC 4045
1883 GCATGAGATCCCGCGCTCGAGGATCATCCAGCCGCGTCCCGGAAAAAGATTCGGAAGC 1824
4046 CCAACCTTTTCATAGAAAGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGGCGTGC 4105
1823 CCAACCTTTTCATAGAAAGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGGCGTGC 1764
4106 CTTGGTGGGTCAFTTTCGAAACCCCGACAGTCCCGCTCAGAAAGAACTCGTCAAGAGCGGATA 4165
1763 CTTGGTGGGTCAFTTTCGAAACCCCGACAGTCCCGCTCAGAAAGAACTCGTCAAGAGCGGATA 1704
4166 GAGGGGATGCGCTCGGAATCGGGAGCGGGATACCGTAAAGCAGCAGGAAAGCGGTACG 4225
1703 GAGGGGATGCGCTCGGAATCGGGAGCGGGATACCGTAAAGCAGCAGGAAAGCGGTACG 1644
4226 CCATTTCGCGCCCAAGCTCTTTCAGCAATATCAAGGTTAGCAAGCTATGTCCTGATAGCG 4285
1643 CCATTTCGCGCCCAAGCTCTTTCAGCAATATCAAGGTTAGCAAGCTATGTCCTGATAGCG 1584
4286 GTCGCGCACACCCAGCGCGGCACAGTCGATGAATCCAGAAAGCGGCAATTTCCACCAT 4345
1583 GTCGCGCACACCCAGCGCGGCACAGTCGATGAATCCAGAAAGCGGCAATTTCCACCAT 1524
4346 GATATTTCGCAAGCAGGCAATCGGCTGAGGTCAGCAGAGATCTCCGCGCTCGGCGCATGCG 4405
1523 GATATTTCGCAAGCAGGCAATCGGCTGAGGTCAGCAGAGATCTCCGCGCTCGGCGCATGCG 1464
4406 CGCCTTAGCGCTGGCGAAAGTTCGGCTGGCGGAGCCCTGATGCTCTTCGTCAGATC 4465
```

Db 1463 GCGCTTGAGCTGCGCAACAGTTTCGGTGGCGGAGCCCTGATGCTCTTCGTCCAGATC 1404
Qy 4466 ATCTGTATCGAAGACCGGCTTCATCCGAGTACGTGCTCGTCTGATGCGATGTTTCGC 4525
Db 1403 ATCTGTATCGAAGACCGGCTTCATCCGAGTACGTGCTCGTCTGATGCGATGTTTCGC 1344
Qy 4526 TTGTGTGCGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGCGGATGTCATCAGC 4585
Db 1343 TTGTGTGCGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGCGGATGTCATCAGC 1284
Qy 4586 CATGATGCGATCTTCTCGGAGGAGCAAGTGGATGACAGAGATCCTGCCCGGAC 4645
Db 1283 CATGATGCGATCTTCTCGGAGGAGCAAGTGGATGACAGAGATCCTGCCCGGAC 1224
Qy 4646 TTGCGCCAAATAGCAGCGATCCCTTCGCTTTCAGTGACAACTGCGAGCAGCTCGCA 4705
Db 1223 TTGCGCCAAATAGCAGCGATCCCTTCGCTTTCAGTGACAACTGCGAGCAGCTCGCA 1164
Qy 4706 AGGAACCGCGTCTGCGGAGGAGCAAGTGGATGACAGAGATCCTGCCCGGAC 4765
Db 1163 AGGAACCGCGTCTGCGGAGGAGCAAGTGGATGACAGAGATCCTGCCCGGAC 1104
Qy 4766 GGCAACCGGAGTGGCTCTTGACAAAGAAACCGGCGCGCTGCGCTGACAGCGGAA 4825
Db 1103 GGCAACCGGAGTGGCTCTTGACAAAGAAACCGGCGCGCTGCGCTGACAGCGGAA 1044
Qy 4826 CAGCGCGCATCAGAGAGCGGATGCTGTTGCGGAGTATGAGCGGATGAGCTCTC 4885
Db 1043 CAGCGCGCATCAGAGAGCGGATGCTGTTGCGGAGTATGAGCGGATGAGCTCTC 984
Qy 4886 CACCAAGCGCGGAGAACTCGGTGCAATCCATCTTGTTCATATGCGAAACGATCC 4945
Db 983 CACCAAGCGCGGAGAACTCGGTGCAATCCATCTTGTTCATATGCGAAACGATCC 924
Qy 4946 TCATCTGCTCTTGATCAGATCTTGATCCCTGCGGCAATCAGATCCTTGCGGCAAGAA 5005
Db 923 TCATCTGCTCTTGATCAGATCTTGATCCCTGCGGCAATCAGATCCTTGCGGCAAGAA 864
Qy 5006 AGCCATCCAGTTACTTTGAGAGGCTTCCCAACTTACAGAGGCGGCGCCAGCTGGCA 5065
Db 863 AGCCATCCAGTTACTTTGAGAGGCTTCCCAACTTACAGAGGCGGCGCCAGCTGGCA 804
Qy 5066 TTCCGTTGCTGCTGCTCATAAAGCGGCGGCTAGCTATGCGCATGTAAAGCCACT 5125
Db 803 TTCCGTTGCTGCTGCTCATAAAGCGGCGGCTAGCTATGCGCATGTAAAGCCACT 744
Qy 5126 GCAAGCTACCTGCTTCTTTCGCTTGGCTTTCCTTGTCCAGATAGCCAGTAGCT 5185
Db 743 GCAAGCTACCTGCTTCTTTCGCTTGGCTTTCCTTGTCCAGATAGCCAGTAGCT 684
Qy 5186 GACATTCATCCGGGTGAGCAGCGTTTCTCGGACTGGCTTTCAGTGTTCGCTTCT 5245
Db 683 GACATTCATCCGGGTGAGCAGCGTTTCTCGGACTGGCTTTCAGTGTTCGCTTCT 624
Qy 5246 TTAGAGCCCTTGGCGGCTGAGCTGCTTGGCGAGCGTG 5283
Db 623 TTAGAGCCCTTGGCGGCTGAGCTGCTTGGCGAGCGTG 586

RESULT 6

LOCUS AY189826/c 2604 bp DNA circular SYN 09-APR-2003
DEFINITION His-3 integration vector pJHAM001, complete sequence.

ACCESSION AY189826

VERSION AY189826.1 GI:28435536

KEYWORDS

SOURCE his-3 integration vector pJHAM001

ORGANISM his-3 integration vector pJHAM001

REFERENCE 1 (bases 1 to 2604)

AUTHORS Lee, D.W., Haag, J.R. and Aramayo, R.

TITLE Construction of strains for rapid homokaryon purification after

integration of constructs at the histidine-3 (his-3) locus of

JOURNAL

PUBMED 12684841

REFERENCE 2 (bases 1 to 2604)

AUTHORS Lee, D.W., Haag, J.R. and Aramayo, R.

TITLE Direct Submission

JOURNAL Submitted (03-DEC-2002) Biology, Texas A&M University, BSBW 415,

College Station, TX 77843-3258, USA

FEATURES Location/Qualifiers

1..2604

source /organism="his-3 integration vector pJHAM001"

/mol_type="genomic DNA"

/db_xref="taxon:220724"

/note="modified pk18 vector"

misc_feature 2542..2572

/note="modified multi-cloning site"

BASE COUNT 561 a 714 c 760 g 569 t

ORIGIN

Query Match 41.2%; Score 2177.4; DB 12; Length 2604;

Best Local Similarity 93.2%; Pred. No. 0;

Matches 2359; Conservative 0; Mismatches 26; Indels 146; Gaps 2;

Qy 2753 CGCTCACAAATCCACACAAATACGAGCGGAGCATAAAGTGTAAAGCCTGGGTGCT 2812

Db 2542 CGCTCACAAATCCACACAAATACGAGCGGAGCATAAAGTGTAAAGCCTGGGTGCT 2485

Qy 2813 AATGAGTGAAGTAACTACATTAATTCGTTGCGTCACTGCCCGCTTTCAGTCGGAA 2872

Db 2484 AATGAGTGAAGTAACTACATTAATTCGTTGCGTCACTGCCCGCTTTCAGTCGGAA 2425

Qy 2873 ACCTGTCGTGCGAGTCATTAATGATCGGCAACGCGGAGGAGCGGTTTGGCTA 2932

Db 2424 ACCTGTCGTGCGAGTCATTAATGATCGGCAACGCGGAGGAGCGGTTTGGCTA 2365

Qy 2933 TTGGCGCTCTTCGCGTCTTCCTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCG 2992

Db 2364 TTGGCGCTCTTCGCGTCTTCCTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCG 2305

Qy 2993 GAGCGGTATCAGTCACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGATACG 3052

Db 2304 GAGCGGTATCAGTCACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGATACG 2245

Qy 3053 CAGGAAAGAAATGTGAGCAAAAGCCAGCAAAAGGCGGAGGAGGCGGCGGCT 3112

Db 2244 CAGGAAAGAAATGTGAGCAAAAGGCGGAGGAGGCGGAGGCGGCGGCT 2185

Qy 3113 TGCTGGCGTCTTTCATAGGCTCCGCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 3172

Db 2184 TGCTGGCGTCTTTCATAGGCTCCGCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 2125

Qy 3173 GTGAGGTGCGGAAACCCGAGCAGACTATAAAGATACAGGCGGTTCCCGCTCGAAGCT 3232

Db 2124 GTGAGGTGCGGAAACCCGAGCAGACTATAAAGATACAGGCGGTTCCCGCTCGAAGCT 2065

Qy 3233 CCCTCGTCCGCTCTCTTTCGCGACCTTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 3292

Db 2064 CCCTCGTCCGCTCTCTTTCGCGACCTTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 2005

Qy 3293 CTTCCGGAAGCTGGCGGCTTCTCATAGCTCACGCTGATGATCTCATGTTCCGTTGAGG 3352

Db 2004 CTTCCGGAAGCTGGCGGCTTCTCATAGCTCACGCTGATGATCTCATGTTCCGTTGAGG 1945

Qy 3353 TCGTTTCGCTCCAAAGTGGGCTGTGTGCAAGAACCCCGCTTTCAGCCCGGCTCGGCT 3412

Db 1944 TCGTTTCGCTCCAAAGTGGGCTGTGTGCAAGAACCCCGCTTTCAGCCCGGCTCGGCT 1885

Qy 3413 TATCCGGAATCTCTTTCGAGTCCCAACCCGCTTTCAGTATGATGATGATGATGATG 3472

Db 1884 TATCCGGAATCTCTTTCGAGTCCCAACCCGCTTTCAGTATGATGATGATGATGATG 1825

Qy 3473 CAGCCACTGGTAAAGGATAGCAGGAGGATGTAGGCGGTGCTACAGATTTCTTGA 3532

Db 3532 CAGCCACTGGTAAAGGATAGCAGGAGGATGTAGGCGGTGCTACAGATTTCTTGA 3472

Db 1824 CAGCCACTGGTAAACAGGATTAGCAGCGAGGATGTAGCGGTCTACAGAGTTCTTGA 1765
QY 3533 AGTGTGGCCTAACTACGGCTACACTAGAGAAAGATATTTGGTATCTGGCCTCTGCTGA 3592
Db 1764 AGTGTGGCCTAACTACGGCTACACTAGAGAAAGATATTTGGTATCTGGCCTCTGCTGA 1705
QY 3593 AGCCAGTTACCTTCGGAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCAGCGTG 3652
Db 1704 AGCCAGTTACCTTCGGAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCAGCGTG 1645
QY 3653 GTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTAACCGCGGAGAAAAAGATCTCAAG 3712
Db 1644 GTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTAACCGCGGAGAAAAAGATCTCAAG 1585
QY 3713 AAGATCCTTTGATCTTTCTACGGGTCTGAGCCTCAGTGAACGAAACTCAAGTTAAG 3772
Db 1584 AAGATCCTTTGATCTTTCTACGGGTCTGAGCCTCAGTGAACGAAACTCAAGTTAAG 1525
QY 3773 GGATTTGTGTCATGAGATTATCGTGCACAAAGCGGCGATCGTGCTCCCACTCCTGCA 3832
Db 1524 GGATTTGTGTCATGAGATTATCAAAAAGGATCTTCACTTAGATCCTT----- 1478
QY 3833 GTTCGGGGCATGGATGCGCGGATAGCGCTGCTGTTTCTCGGATGCGAGGATTTGC 3892
Db 1477 ----- 1478
QY 3893 ACTGCGGTAGAACTCCGCGAGTGCTCCAGCCTCAGCGAGCTGAACCAACTCGCGA 3952
Db 1477 ----- 1478
QY 3953 GGGGATCGAGCCCGGGTGGCGAGAACTCCAGCATGAGATCCCGCGCTGAGGATCA 4012
Db 1477 -----TTGGGGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGAGGATCA 1429
QY 4013 TCCAGCGCGGTCCCGGAAAAAGATTCCGAAGCCCAACCTTTATAGAGAGCGCGGTGG 4072
Db 1428 TCCAGCGCGGTCCCGGAAAAAGATTCCGAAGCCCAACCTTTATAGAGAGCGCGGTGG 1369
QY 4073 ANTGAAATCTGATGAGGAGTTGGGGTTCGCTTGGTGGCTCATTTGACACCCAGAG 4132
Db 1368 AATCGAAATCTGATGAGGAGTTGGGGTTCGCTTGGTGGCTCATTTGACACCCAGAG 1309
QY 4133 TCCCGCTCAGAGAACTCGTCAAGAGGGGATAGAGAGGATGCGTGCAGATTCGGGAGC 4192
Db 1308 TCCCGCTCAGAGAACTCGTCAAGAGGGGATAGAGAGGATGCGTGCAGATTCGGGAGC 1249
QY 4193 GGGGATACCGTAAAGCAGAGGAGGCTCAGCCCATTTGCGCGCAAGCTTTCAGCAAT 4252
Db 1248 GGGGATACCGTAAAGCAGAGGAGGCTCAGCCCATTTGCGCGCAAGCTTTCAGCAAT 1189
QY 4253 ATCAGGGTAGCCAAAGCTATGCTGATAGCGGTCCGCCACACCCAGCGCGCCAGTGC 4312
Db 1188 ATCAGGGTAGCCAAAGCTATGCTGATAGCGGTCCGCCACACCCAGCGCGCCAGTGC 1129
QY 4313 GATGAATCCAGAAAGCGGCATTTTCCACATGATATTCGGCAAGCAGCATTCGCCATG 4372
Db 1128 GATGAATCCAGAAAGCGGCATTTTCCACATGATATTCGGCAAGCAGCATTCGCCATG 1069
QY 4373 GGTCAAGCAGATCTTCGGCTCGGATCGGCGCTTCCAGCTGCGCAACAGTTTCGCG 4432
Db 1068 GGTCAAGCAGATCTTCGGCTCGGATCGGCGCTTCCAGCTGCGCAACAGTTTCGCG 1009
QY 4433 TGGCGGAGCCCTGATGCTTTCGTCAGATCATCTGATCGACAGACCGGCTTCCAT 4492
Db 1008 TGGCGGAGCCCTGATGCTTTCGTCAGATCATCTGATCGACAGACCGGCTTCCAT 949
QY 4493 CCGAGTAGCTCTCGCTCGATGCTTTCGCTTGGTGGTCCGATGGGCGAGTAGCCG 4552
Db 948 CCGAGTAGCTCTCGCTCGATGCTTTCGCTTGGTGGTCCGATGGGCGAGTAGCCG 889
QY 4553 ATCAAGCGTATCGAGCGCGCATTTGATCAGCCATGATGATGATGATGATGATGATG 4612
Db 888 ATCAAGCGTATCGAGCGCGCATTTGATCAGCCATGATGATGATGATGATGATGATG 829

QY 4613 AAGGTGAGATGACAGGAGATCTGCGCCCGCACTTTCGCCCAATAGCAGCAGTCCCTTCC 4672
Db 828 AAGGTGAGATGACAGGAGATCTGCGCCCGCACTTTCGCCCAATAGCAGCAGTCCCTTCC 769
QY 4673 CGTTTCAGTGACAAAGTCGAGCAGCTGCGCAAGAGAGCGCGTGGTGGCGAGCAGCA 4732
Db 768 CGTTTCAGTGACAAAGTCGAGCAGCTGCGCAAGAGAGCGCGTGGTGGCGAGCAGCA 709
QY 4733 TAGCCCGCTGCTGCTGCTGCTGAGTTTCAATTCAGGCGCACCGCAGAGCTGCTTGA 4792
Db 708 TAGCCCGCTGCTGCTGCTGAGTTTCAATTCAGGCGCACCGCAGAGCTGCTTGA 649
QY 4793 AAGAAACGGGCGGCTGCTGCTGAGCAGCGGAGACACGGCGGATCAGACAGCGGATCT 4852
Db 648 AAGAAACGGGCGGCTGCTGCTGAGCAGCGGAGACACGGCGGATCAGACAGCGGATCT 589
QY 4853 CTGTTGTGCCAGTATAGCCGAATAGCTCTCCACCCAGCGCGGAGAACCTCGGTG 4912
Db 588 CTGTTGTGCCAGTATAGCCGAATAGCTCTCCACCCAGCGCGGAGAACCTCGGTG 529
QY 4913 CAATCCATCTTTGTTCAATCATGCGAAACGATCTCTCTCTGATCAGATCTTGA 4972
Db 528 CAATCCATCTTTGTTCAATCATGCGAAACGATCTCTCTCTGATCAGATCTTGA 469
QY 4973 TCCCTCGGCGCATCAGATCTTGGCGGCAAGAAAGCAATCCAGTTTACCTTTCAGGCTT 5032
Db 468 TCCCTCGGCGCATCAGATCTTGGCGGCAAGAAAGCAATCCAGTTTACCTTTCAGGCTT 409
QY 5033 CCCAACCTTACCAGAGGGCGCCCAAGTCCGTTTCGCTTCTGCTGCTTCCATTAAC 5092
Db 408 CCCAACCTTACCAGAGGGCGCCCAAGTCCGTTTCGCTTCTGCTGCTTCCATTAAC 349
QY 5093 CGCCAGCTAGCTATGCGCATCTAGCCCACTGAGCTTACCTGCTTCTTCTTTCGCT 5152
Db 348 CGCCAGCTAGCTATGCGCATCTAGCCCACTGAGCTTACCTGCTTCTTCTTTCGCT 289
QY 5153 TGCCTTTTCCCTTCCAGATAGCCCACTGAGCTTCAATCCGCGGTGAGCACCGTTT 5212
Db 288 TGCCTTTTCCCTTCCAGATAGCCCACTGAGCTTCAATCCGCGGTGAGCACCGTTT 229
QY 5213 CTGCGGAGCTGCTTCTGCTTTCGCTTTCCTTTCAGCCCTTTCGCGCTCAGTGCCT 5272
Db 228 CTGCGGAGCTGCTTCTGCTTTCGCTTTCCTTTCAGCCCTTTCGCGCTCAGTGCCT 169
QY 5273 GCGGCGAGCTG 5283
Db 168 GCGGCGAGCTG 158

RESULT 7
AY189827/cLOCUS 11373 bp DNA circular SYN 09-APR-2003
DEFINITION His-3 integration vector pJHAM002, complete sequence.

ACCESSION AY189827

VERSION AY189827.1 GI:28435537

KEYWORDS

SOURCE his-3 integration vector pJHAM002

ORGANISM

his-3 integration vector pJHAM002

REFERENCE

1 (bases 1 to 11373)

AUTHORS

Lee, D.W., Haag, J.R., and Aramayo, R.

TITLE

Construction of strains for rapid homokaryon purification after

integration of constructs at the histidine-3 (his-3) locus of

Neurospora crassa

Curr. Genet. 43 (1), 17-23 (2003)

12684841

JOURNAL

2 (bases 1 to 11373)

PUBMED

Lee, D.W., Haag, J.R., and Aramayo, R.

AUTHORS

Direct Submission

TITLE

Submitted (03-DEC-2002) Biology, Texas A&M University, BSBW 415,

College Station, TX 77843-3258, USA

JOURNAL

Location/Qualifiers

FEATURES


```
Db 860 AAGGTGAGATGACAGGAGATCCTGCCCGGACATTCGCCCAATAGACGACGATCCCTTCC 801
Qy 4673 CGCTTCAGTGAACAAGTGCAGACAGTGCAGCAAGGAAACGCCCTCGTGGCCAGCCACGA 4732
Db 800 CGCTTCAGTGAACAAGTGCAGACAGTGCAGCAAGGAAAGGCCCTCGTGGCCAGCCACGA 741
Qy 4733 TAGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4792
Db 740 TAGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
Qy 4793 AAGAACCGGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4852
Db 680 AAGAACCGGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
Qy 4853 CTGTTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4912
Db 620 CTGTTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
Qy 4913 CAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4972
Db 560 CAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
Qy 4973 TCCCTGCGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5032
Db 500 TCCCTGCGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
Qy 5033 CCCAACCTTACAGAGGGGCGCCCGAGCTGGCAATTCGGTTCGCTGCTGCTGCTGCTGCTGCTGCT 5092
Db 440 CCCAACCTTACAGAGGGGCGCCCGAGCTGGCAATTCGGTTCGCTGCTGCTGCTGCTGCTGCTGCT 381
Qy 5093 CGCCGAGCTAGCTATGCGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5152
Db 380 CGCCGAGCTAGCTATGCGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
Qy 5153 TGGGTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5212
Db 320 TGGGTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 261
Qy 5213 CTGCGGAGCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5272
Db 260 CTGCGGAGCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
Qy 5273 GCGGACAGGTG 5283
Db 200 GCGGACAGGTG 190
```

```
RESULT 8
LOCUS AY189829/c 11403 bp DNA circular SYN 09-APR-2003
DEFINITION His-3 integration vector pUHAM004, complete sequence.
ACCESSION AY189829
VERSION AY189829.1 GI:28435539
KEYWORDS his-3 integration vector pUHAM004
SOURCE his-3 integration vector pUHAM004
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 11403)
AUTHORS Lee,D.W., Haag,J.R. and Aramayo,R.
TITLE Construction of strains for rapid homokaryon purification after
integration of constructs at the histidine-3 (his-3) locus of
Neurospora crassa
JOURNAL Curr. Genet. 43 (1), 17-23 (2003)
PUBMED 12684841
REFERENCE 2 (bases 1 to 11403)
AUTHORS Lee,D.W., Haag,J.R. and Aramayo,R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2002) Biology, Texas A&M University, BSW 415,
College Station, TX 77843-3258, USA
FEATURES
LOCATION/Qualifiers
source 1..11403
/organism="his-3 integration vector pUHAM004"
/mol_type="genomic DNA"
```

```
/db_xref="taxon:220727"
/focus 1..2597
/organism="his-3 integration vector pUHAM001"
/mol_type="genomic DNA"
/db_xref="taxon:220724"
misc_feature 2598..11403
/feature="his-3 HindIII fragment"
misc_feature 7793..7845
/feature="multi-cloning sites"
BASE COUNT 2554 a 3059 c 3033 g 2757 t
ORIGIN
Query Match 41.2%; Score 2177.4; DB 12; Length 11403;
Best Local Similarity 93.2%; Pred No. 0;
Matches 2359; Conservative 0; Mismatches 26; Indels 146; Gaps 2;
Qy 2753 CGCTCAAAATCCACAAACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTT 2812
Db 2574 CGCTCAAAATCCACAAACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTT 2517
Qy 2813 AATGAGTGAGCTAACTCACTAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2872
Db 2516 AATGAGTGAGCTAACTCACTAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2457
Qy 2873 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2932
Db 2456 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2397
Qy 2933 TTGGGCGCTTCTCCGCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2992
Db 2396 TTGGGCGCTTCTCCGCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2337
Qy 2993 GAGCGGTATCAGCTCACTCAAAAGCGGTAAATAGGTTATCCACAGAAATCAGGGGATAACG 3052
Db 2336 GAGCGGTATCAGCTCACTCAAAAGCGGTAAATAGGTTATCCACAGAAATCAGGGGATAACG 2277
Qy 3053 CAGGAAAGAAACATGTGAGCAAAAGGCGGCAAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 3112
Db 2276 CAGGAAAGAAACATGTGAGCAAAAGGCGGCAAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2217
Qy 3113 TGCTGGGCTTTTCCATAGGCTCCGCCCGCTGACGAGCATCACAAAATTCAGCGCTCAA 3172
Db 2216 TGCTGGGCTTTTCCATAGGCTCCGCCCGCTGACGAGCATCACAAAATTCAGCGCTCAA 2157
Qy 3173 GTCAGAGGTGGGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTTGGAGGCT 3232
Db 2156 GTCAGAGGTGGGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTTGGAGGCT 2097
Qy 3233 CCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3292
Db 2096 CCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2037
Qy 3293 CTTCCGGAAGCGTGGGCTTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3352
Db 2036 CTTCCGGAAGCGTGGGCTTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1977
Qy 3353 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3412
Db 1976 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1917
Qy 3413 TATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3472
Db 1916 TATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1857
Qy 3473 CAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3532
Db 1856 CAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797
Qy 3533 AGTGGTGGCTTAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3592
Db 1796 AGTGGTGGCTTAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1737
```


2316	TCCTTGAGTCCAA	CCCGGTAA	CA	CGACTTATCG	CGCACTGG	CAGCAG	CA	CGTAA	CAG	2325
3489	GATTAGCAGACG	GAGGTATG	TAGGCGGTG	CTACAGAGTTC	TTTGAAGTGG	TGGTGG	CTTAA	CTA	3548	
2256	GATTAGCAGACG	GAGGTATG	TAGGCGGTG	CTACAGAGTTC	TTTGAAGTGG	TGGTGG	CTTAA	CTA	2197	
3549	CGGCTACACTA	GAGAACA	CAGTATTT	TGGTATCT	CGGCTCTG	CTGCTGA	AGCCAG	TACTCTCGG	3608	
2196	CGGCTACACTA	GAGGACAG	TATTTGGTATCT	CGGCTCTG	CTGCTGA	AGCCAG	TACTCTCGG	2137		
3609	AAAAAGATTTG	TGATCTTT	GATCCGGCAAA	CAAACCA	CCGCTGGT	TAGCGGTGGT	TTTTT	3658		
2136	AAAAAGATTTG	TGATCTTT	GATCCGGCAAA	CAAACCA	CCGCTGGT	TAGCGGTGGT	TTTTT	2077		
3669	TGCTTTGCAAG	CAGCAGAT	TACGCGCAG	AAAAAAGG	ATCTCA	AGAAGATCT	TTTGTGATCTT	3728		
2076	TGCTTTGCAAG	CAGCAGAT	TACGCGCAG	AAAAAAGG	ATCTCA	AGAAGATCT	TTTGTGATCTT	2017		
3729	TTCTACGGGGT	CTGACGCT	CAGTGGAA	CGAAAACT	CACGTTA	AGGATTTTT	TGGTCATGAG	3788		
2016	TTCTACGGGGT	CTGACGCT	CAGTGGAA	CGAAAACT	CACGTTA	AGGATTTTT	TGGTCATGAG	1957		
3789	ATTATCGTGG	-----	-----	-----	-----	-----	-----	3798		
1956	TTTTCGCTGC	ATAACCGT	CTGCGGGT	CAATATAG	CGATTTTT	TCGGTATAT	TCCATCC	1897		
3799	-----	-----	-----	-----	-----	-----	-----	3798		
1896	TTTTTCGCAG	TATACAGGA	TTTTTGC	CAAAGGTT	TCGTGTAG	ACTTTCT	TGGTGTATC	1837		
3799	-----	-----	-----	-----	-----	-----	-----	3840		
1836	CAACGGCGT	CAGCGGGC	GAGATAG	TAGGCGCA	CCCGC	GAGCGGGT	TTCTTTC	1777		
3841	GCATGATG	CGCGGAT	AGCGCTCT	CTGGTTCT	CTGATCG	CAACGGA	TTGCACTG	3900		
1776	TTCACTGT	CCCTTATT	CGCACTGG	CGGTGCT	CAA	CGGAA	TCTTCTCT	1717		
3901	TAGAACTCCG	-----	-----	-----	-----	-----	-----	3947		
1716	CGGGCTAC	CGCGCGTAA	CAGATGAG	GGCAAGCG	GATGGCTG	ATGA	AAACCAAG	1657		
3948	CGCGAGGGAT	TCGAGCCG	GGGTGGGCG	AGAAATCC	CACGATG	AGATCCC	CGCGCTGAG	4007		
1656	AGGAAGGG	CAGGCC	ACCTATCA	AGGTGTAT	TGCTTCC	AGACGA	AGACGATGAG	1597		
4008	GATCATCCAG	-----	-----	-----	-----	-----	-----	4018		
1596	GAAAGGCG	GGCGGGC	GCATGAG	CTTCTG	CGCTTAC	CTGTG	CGCGTCCG	1537		
4019	-----	-----	-----	-----	-----	-----	-----	4037		
1536	TACAAATCA	CGGGCGT	GTGGACTAT	GACACGT	CCGAGGG	CGTCCG	GAAACGAT	1477		
4038	TCCGAGCCCA	CTTTCATAG	AGGGCGGGT	TGGAATCG	AAATCT	CTGTGAT	TGGCAGGTT	4097		
1476	TCCGAGCCCA	CTTTCATAG	AGGGCGGGT	TGGAATCG	AAATCT	CTGTGAT	TGGCAGGTT	1417		
4098	GGCGTCTG	TGTCGTCT	TTTGA	ACCCAG	AGTCCG	CTCAG	AAAGAACTCT	4157		
1416	GGCGTCTG	TGTCGTCT	TTTGA	ACCCAG	AGTCCG	CTCAG	AAAGAACTCT	1357		
4158	AGCGATAG	AGCGAT	CGCTGGA	ATCCGG	AGCGCG	CGATAC	CGTAAAGCAG	4217		
1356	AGCGATAG	AGCGAT	CGCTGGA	ATCCGG	AGCGCG	CGATAC	CGTAAAGCAG	1297		
4218	CGGTAC	GGCCATCT	TCCGCC	CAAGCTT	TTTCA	GCAATAT	TCACGGGT	4277		
1296	CGGTAC	GGCCATCT	TCCGCC	CAAGCTT	TTTCA	GCAATAT	TCACGGGT	1237		
4278	TGATACGG	TCCGCAC	ACCCAG	CGGCCAC	AGCTG	ATGATCC	AGAAAGCG	4337		
1236	TGATACGG	TCCGCAC	ACCCAG	CGGCCAC	AGCTG	ATGATCC	AGAAAGCG	1177		

```
QY 4338 TCCACATGATATTCGGACAGAGGATCGCCATCGGTCACGACGATCCTCGCGTCG 4397
Db 1176 TCCACATGATATTCGGACAGAGGATCGCCATCGGTCACGACGATCCTCGCGTCG 1117
QY 4398 GGCATCGCGCCCTTGAGCTCGCGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCG 4457
Db 1116 GGCATCGCGCCCTTGAGCTCGCGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCG 1057
QY 4458 TCCAGATCATCTGATCGACAGACCGGGTTCATCCGAGTAGCTGCTGCTCGATGCGA 4517
Db 1056 TCCAGATCATCTGATCGACAGACCGGGTTCATCCGAGTAGCTGCTGCTCGATGCGA 997
QY 4518 TCTTTTCGTTGCTGCTGATCGGAGTTCAGGATCAAGGATGATGAGCGCGCGGAT 4577
Db 996 TCTTTTCGTTGCTGCTGATCGGAGTTCAGGATCAAGGATGATGAGCGCGCGGAT 937
QY 4578 GCATCAGCCATGATGATATCTTCTCGCAGGAGCAAGGTGAGATGACAGGATCCTGC 4637
Db 936 GCATCAGCCATGATGATATCTTCTCGCAGGAGCAAGGTGAGATGACAGGATCCTGC 877
QY 4638 CCGGCACTTCGCCAATAGCAGCAGTCCCTTCGCGTTCAGTGACAACTCGAGCACA 4697
Db 876 CCGGCACTTCGCCAATAGCAGCAGTCCCTTCGCGTTCAGTGACAACTCGAGCACA 817
QY 4698 GCTCGCAAGGAACGCCGCTGCTGCGCCAGCAGATAGCCGCTGCTGCTCGTCACT 4757
Db 816 GCTCGCAAGGAACGCCGCTGCTGCGCCAGCAGATAGCCGCTGCTGCTCGTCACT 757
QY 4758 TCATTCAGGACACCGGACAGTCTGTTGACAAAAGAACCGGGCGCCCTCGCTGAC 4817
Db 756 TCATTCAGGACACCGGACAGTCTGTTGACAAAAGAACCGGGCGCCCTCGCTGAC 697
QY 4818 AGCGGGAACACCGGGGATCAGACAGCCGATTCGTTGTTGTCGCCAGTCATGCGCAAT 4877
Db 696 AGCGGGAACACCGGGGATCAGACAGCCGATTCGTTGTTGTCGCCAGTCATGCGCAAT 637
QY 4878 AGCTCTCCACCCAGCGGCGGAGAACCTGCGTGCAATCCATCTGTTCAATCATGCGA 4937
Db 636 AGCTCTCCACCCAGCGGCGGAGAACCTGCGTGCAATCCATCTGTTCAATCATGCGA 577
QY 4938 AACGATCTCATCTGCTCTTGTATCAGATCTGATCCCTGCGGCAATCATGCTTGGC 4997
Db 576 AACGATCTCATCTGCTCTTGTATCAGATCTGATCCCTGCGGCAATCATGCTTGGC 517
QY 4998 GGCAGGAAGCCATCCAGTTTACTTTTCAGGGCTTCCCAACCTTACAGAGGCGGCCCA 5057
Db 516 GGCAGGAAGCCATCCAGTTTACTTTTCAGGGCTTCCCAACCTTACAGAGGCGGCCCA 457
QY 5058 GCTGGCAATTCGGTTGCTGTCATATAAACCGCCAGCTAGCTATGCGCATGTA 5117
Db 456 GCTGGCAATTCGGTTGCTGTCATATAAACCGCCAGCTAGCTATGCGCATGTA 397
QY 5118 AGCCCACTGCAAGTACTGCTTCTCTTTGCGCTGCTGCTTTTCCCTTGCAGATAGCC 5177
Db 396 AGCCCACTGCAAGTACTGCTTCTCTTTGCGCTGCTGCTTTTCCCTTGCAGATAGCC 337
QY 5178 CAGTAGTGATTCATCCGGGTGACACCGTTTTCGCGAATGCTGTTTCTACGTGTTTC 5237
Db 336 CAGTAGTGATTCATTCAGCGGTGACACCGTTTTCGCGAATGCTGTTTCTACGTGTTTC 277
QY 5238 CGCTTCTTTAGCAGCCCTTGGCCCTGAGTGTGCTGCGGAGCGTG 5283
Db 276 CGCTTCTTTAGCAGCCCTTGGCCCTGAGTGTGCTGCGGAGCGTG 231
```

```
RESULT 10
AY219859/c
LOCUS AY219859 3210 bp DNA circular SYN 23-MAR-2003
DEFINITION Cloning vector pK19MECA2, complete sequence.
ACCESSION AY219859
VERSION AY219859.1 GI:29164979
KEYWORDS
```

```
SOURCE Cloning vector pK19MECA2
ORGANISM Cloning vector pK19MECA2
REFERENCE 1 (bases 1 to 3210)
AUTHORS Kirchner, O. and Tauch, A.
TITLE Tools for genetic engineering in the amino acid-producing bacterium
Corynebacterium glutamicum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3210)
AUTHORS Kirchner, O. and Tauch, A.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2003) Department of Genetics, University of
Bielefeld, Universitaetstrasse 25, Bielefeld D-33615, Germany
FEATURES
    Location/Qualifiers
        1..3210
            /organism="Cloning vector pK19MECA2"
            /mol_type="genomic DNA"
            /db_xref="taxon:224628"
            755..1549
            /gene="aph(3')-IIa"
            755..1549
            /gene="aph(3')-IIa"
            /note="aminoglycoside phosphotransferase"
            /codon_start=1
            /transl_table=11
            /product="aph(3')-IIa"
            /protein_id="AA065226.1"
            /db_xref="GI:29164980"
            /translation="MIEQDLHAGSPAAWVERLFQYDWAQTTIGSDAAVRLSQAQR
            PVLFTDLISALNELODEARLTLWATTPGCAAVLDVTEAGRDWLLGLGVQDIL
            LSHLPAEKVSIIMADAMRLHTLDPATCPFDHQAHRIRERATMEAGLVDDDLDE
            BHOGLAPAEFLPARLKAMPDGEDLVVTHGDACLNIWVENGFRSGFIDCGLGVADRY
            ODIALATRDIAELGGEWADFLVLYGAPDSORIAFRLULDEFF"
            1934..2032
            /misc_feature
            /note="RP4-mob region"
            2795..2797
            /rep_origin
            /note="oriV"
BASE COUNT 705 a 887 c 919 g 699 t
ORIGIN
Query Match 39.5%; Score 2088.4; DB 12; Length 3210;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 141; Indels 231; Gaps 3;
QY 2709 GACGTAATCATGTCATAGCTGTTTCTGTCGTGAAATGTTATCGCTCACAATTCACA 2768
Db 3210 GCGTAATCATGTCATAGCTGTTTCTGTCGTGAAATGTTATCGCTCACAATTCACA 3151
QY 2769 CAACATACGAGCGGAGCATATAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAT 2828
Db 3150 CAACATACGAGCGGAGCATATAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAT 3091
QY 2829 CACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTCGGGAAACCTGTCGCCAGCT 2888
Db 3090 CACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTCGGGAAACCTGTCGCCAGCT 3031
QY 2889 GCATTAATGAATCGCCCAACCGCGGGGAGAGCGGGTTTGGCTATTTGGCGGCTCTTCGCG 2948
Db 3030 GCATTAATGAATCGCCCAACCGCGGGGAGAGCGGGTTTGGCTATTTGGCGGCTCTTCGCG 2971
QY 2949 TTCCTCGCTCACTGACTGCTGCGCTCGTTCGCTGCGGTGCGGAGCGGTATCAGTCA 3008
Db 2970 TTCCTCGCTCACTGACTGCTGCGCTCGTTCGCTGCGGTGCGGAGCGGTATCAGTCA 2911
QY 3009 CTCAAAGCGCGTAAATACGTTTATCCAGAAATCAGGGGATAACCGAGGAAGACATGTG 3068
Db 2910 CTCAAAGCGCGTAAATACGTTTATCCAGAAATCAGGGGATAACCGAGGAAGACATGTG 2851
QY 3069 AGCAAAAGCGCAGCAAAAGCGCAGAACCGGTAAAGAGCGCGGTGCTGCGGTTTTTCCA 3128
Db 2850 AGCAAAAGCGCAGCAAAAGCGCAGAACCGGTAAAGAGCGCGGTGCTGCGGTTTTTCCA 2791
QY 3129 TAGGCTCGCGCCCTTGCAGGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 3188
```



```
QY 5118 ACCCATGCTAGAGTACCTGCTTCTCTTTGGCTGCTGCTTTCCCTTGTCCAGATAGCC 5177
Db 570 AGCCCACTGCAAGTACCTGCTTCTCTTTGGCTGCTGCTTTCCCTTGTCCAGATAGCC 511
QY 5178 CAGTAGCTGACATTCATCCGGGCTCAGCACCGTTTCTCGGACTGGCTTCTACGTGTTCC 5237
Db 510 CAGTAGCTGACATTCATCCGGGCTCAGCACCGTTTCTCGGACTGGCTTCTACGTGTTCC 451
QY 5238 CGCTTCCTTTAGACGCCCTTGGCCCTCAGTGTCTTGGCGACGCTG 5283
Db 450 CGCTTCCTTTAGACGCCCTTGGCCCTCAGTGTCTTGGCGACGCTG 405

RESULT 11
AY228222/c
LOCUS      5695 bp      DNA      circular SYN 25-WAR-2003
DEFINITION      Shuttle vector pEC-K19mob2, complete sequence.
ACCESSION      AY228222
VERSION        AY228222.1 GI:29242905
KEYWORDS
SOURCE
ORGANISM      Shuttle vector pEC-K19mob2
                Shuttle vector pEC-K19mob2
                artificial sequences; vectors.
REFERENCE      1 (bases 1 to 5695)
                Kirchner,O. and Tauch,A.
                Tools for genetic engineering in the amino acid-producing bacterium
                Corynebacterium glutamicum
                Unpublished
                2 (bases 1 to 5695)
                Kirchner,O. and Tauch,A.
                Direct Submission
                Submitted (22-JAN-2003) Department of Genetics, University of
                Bielefeld, Universitaetsstrasse 25, Bielefeld D-33615, Germany
FEATURES
    source
        1..5695
            /organism="Shuttle vector pEC-K19mob2"
            /mol_type="genomic DNA"
            /db_xref="taxon:224992"
            /note="Escherichia coli-Corynebacterium glutamicum shuttle
            vector"
        400..1863
            /gene="repa"
        400..1863
            /gene="repa"
        /note="replication protein"
        /codon_start=1
        /transl_table=11
        /product="Repa"
        /protein_id="AA066604.1"
        /db_xref="GI:29242906"
        /translation="MTLADPDQTVTASAWKLSADLPDTHPEAMRCGSRGWTADREEL
        LAHGRSFGSKTRDFASAWIKNPDTGETQPKLYRAGSKALTRCOYVALTHAQHAHV
        IVLIDYPSHQAGKIEHVNPOYVAILLEKWARLEKAPWIGNPLSGKQLHILDPV
        YAAKGTSPNRLIAATEEMRTVFGADQAFSHLSRWPLHVSDDPTAYKWHQCHDRV
        DLADLMEIARTWTSQPKKYIEQDFSSGRARI EAAQRATAEKALALIDASLPAL
        DASGLIDGVYRWLTNPRARDETAFPHALTGVOLKAGERLKDAKALIDAVEAYNV
        AQAAGADGREFDLPAMDRLTMAARVRGVYVAKGPVVPARVETQSSRGKALATWGR
        RGAATSNARWADPESKYAQETRORLAEANKREMTGELLELVKLTALIDARSQSVAD
        PSTRELAGELVGSERRIQVYRKALGMEAKGRPRAEN"
        2506..2721
            /gene="per"
        2506..2721
            /gene="per"
        /note="positive effector of replication"
        /codon_start=1
        /transl_table=11
        /product="per"
        /protein_id="AA066605.1"
        /db_xref="GI:29242907"
        /translation="MDLSVDVALESALDAVLDLKTIVGFFRALDITDAPASRDWASAA
        SDLETLVADLELADELARQRQEDAQ"
        3240..4034
            /gene="aph(3')-IIa"
            /db_xref="GI:29242908"
            /translation="MIEODGLHAGSPAAWVERLFGVDWAQOTIGSDAAVPELSAGGR
            PVLVFKTDLSGALNELQDEARLSWLTGVCFAVLVDTVEAGRDWLLGLSVFGDDL
            LSHLAPAEKVSIMADAMRRLHLDPATCFPDHQAKHRIERARTMEAGLVDDDLDE
            EHQLAPAEFLFARKRMPDGEDLVTHGDACLPLNINVENGSRFSFGIDGRLGVADRY
            ODIALATRDIAEELGGEWADRLVLYGIAAPDSQRIAFYRLIDEPF"
            4419..4517
                /misc_feature
            /note="RP4-mob region"
            5280..5282
                /rep_origin
            /note="oriV"
BASE COUNT      1285 a      1626 c      1636 g      1148 t
ORIGIN
```

```
QY 2709 GACGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACAA 2768
Db 5695 GCGGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACAA 5636
QY 2769 CAACATACGAGCGGGAAGCATAAAGTGTAAAGCCGCGGTGCTTAATCAGTGAGCTAACT 2828
Db 5635 CAACATACGAGCGGGAAGCATAAAGTGTAAAGCCGCGGTGCTTAATCAGTGAGCTAACT 5576
QY 2829 CACATTAATTCGTTGCGCTCACTCCCGCTTTCAGTTCGGGAAACCTGTCTGTCAGCT 2888
Db 5575 CACATTAATTCGTTGCGCTCACTCCCGCTTTCAGTTCGGGAAACCTGTCTGTCAGCT 5516
QY 2889 GCATTAATTAATTCGCGCCACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCG 2948
Db 5515 GCATTAATTAATTCGCGCCACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCG 5456
QY 2949 TTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGCGGAGCGGTATCAGCTCA 3008
Db 5455 TTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCGGCGGAGCGGTATCAGCTCA 5396
QY 3009 CTCAGAGCGGTATACAGGTTATCCACAGATCAGGGGATTAACGAGGAAAGCAATGTG 3068
Db 5395 CTCAGAGCGGTATACAGGTTATCCACAGATCAGGGGATTAACGAGGAAAGCAATGTG 5336
QY 3069 AGCAAAAGCGCCAGCAAAAGGCGCAGGAACCGTAAAGAGCGCGGTTCGTCGCGTTCCTCCA 3128
Db 5335 AGCAAAAGCGCCAGCAAAAGGCGCAGGAACCGTAAAGAGCGCGGTTCGTCGCGTTCCTCCA 5276
QY 3129 TAGGTCGCGCGCGCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCGCGAAA 3188
Db 5275 TAGGTCGCGCGCGCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGCGCGAAA 5216
QY 3189 CCGGACAGGCTATTAAGATACACAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCGCTTCCC 3248
Db 5215 CCGGACAGGCTATTAAGATACACAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCGCTTCCC 5156
QY 3249 TGTTCCGACCTTCGCGCTTACCGGATACCTGTCTCGCGCTTTTCTCCCTTCGGGAAGCGTGGC 3308
Db 5155 TGTTCCGACCTTCGCGCTTACCGGATACCTGTCTCGCGCTTTTCTCCCTTCGGGAAGCGTGGC 5096
QY 3309 GCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGCTGAGTTCGCTTCGCTCAAGCT 3368
Db 5095 GCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGCTGAGTTCGCTTCGCTCAAGCT 5036
QY 3369 GGGCTGTGTGACGAACCCCGCTTCAGCCGACGCTTCGCGCTTATCCCGTAACTATCG 3428
Db 5035 GGGCTGTGTGACGAACCCCGCTTCAGCCGACGCTTCGCGCTTATCCCGTAACTATCG 4976
```

3429 TCTTGAGTCAACCCCGTAAAGACAGACTTATCGCAGCTGGCAGCAGCCACTGTAACAG 3488
Db TCTTGAGTCAACCCCGTAAAGACAGACTTATCGCAGCTGGCAGCAGCCACTGTAACAG 4916
3489 GATTACAGAGCAGAGTATGTAAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTA 3548
Db GATTACAGAGCAGAGTATGTAAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTA 4856
3549 CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTACCTTCGG 3608
Db CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTACCTTCGG 4796
3609 AAAAAGAGTGTAGTCTTGATCGGCAAAACAAACCCAGCTGCTAGCGGTGGTTTTTT 3668
Db AAAAAGAGTGTAGTCTTGATCGGCAAAACAAACCCAGCTGCTAGCGGTGGTTTTTT 4736
3669 TGTTCGAAGCAGCAGTACGCGCAGAAAAAAGATCTCAAGAAAGATCTCTTTGATCTT 3728
Db TGTTCGAAGCAGCAGTACGCGCAGAAAAAAGATCTCAAGAAAGATCTCTTTGATCTT 4676
3729 TTCTACGGGTCTGAGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTGGTCAATGAG 3788
Db TTCTACGGGTCTGAGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTGGTCAATGAG 4616
3789 ATTATCGTCG----- 3798
Db ATTATCGTCG----- 4556
3799 ----- 3798
Db ----- 4496
4555 TTTTTCGCACGATATACAGGATTTTGCACAAAGGGTTCGTGTAGACTTTCTTGGTGTATC 4496
3799 ----- 3840
Db ----- 4436
4495 CAACGGCTCAGCGGGCAGGATAGTGAAGTAGGCGCACCCCGCAGCGGGTTCCTTC 3900
3841 GCATGGATCGCGGATAGCGCTGCTGTTTCTCGATGCCAGCGATTTGCACTGCCGG 4436
Db TTTCACTGTCCTTATTCGCACCTGGCGGTCTCAACGGGAATCTGCTCTCGAGGCTGG 4376
3901 TAGAATCTCGC----- 3947
Db ----- 4316
4375 CCGGCTACCGCGCGGCTACAGATGAGGCAAGCGGATGGCTGTATGAAACCAAGCAACC 4007
3948 CGGAGGGGATCGAGCCCGGGTGGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAG 4256
Db AGGAAGGCGAGCCCACTATCAGGTGTACTGCTTCAGACGACGAGAGCGATTCAG 4018
4008 GATCATCTCAGC----- 4196
Db GAAAGCGCGCGCGCGCGCATAGCCTGTGCGGCTACCTGCTGGCGGTGGCCAGGCG 4037
4019 ----- 4136
Db TACAAAATCAGGGCGTGTGACTATGACAGCTGCGGAGGGCGTCCCGGAAACGAT 4097
4038 TCCGAAGCCAACTTTCTAGAGCGCGGTGGAAATCGAAATCTCGTGTAGCGAGTT 4135
Db TCCGAAGCCAACTTTCTAGAGCGCGGTGGAAATCGAAATCTCGTGTAGCGAGTT 4076
4098 GGGCGTCTGCTGCTGCTATTCGAAACCCAGAGTCCCGCTCAGAAACTCTGTCAAGA 4157
Db GGGCGTCTGCTGCTGCTATTCGAAACCCAGAGTCCCGCTCAGAAACTCTGTCAAGA 4016
4075 GGGCGTCTGCTGCTGCTATTCGAAACCCAGAGTCCCGCTCAGAAACTCTGTCAAGA 4217
4158 AGGCGATAGAAGCGGATGGCTGCGAATCGGAGCGCGGATACCGTAAAGCACAGGAG 4015
Db AGGCGATAGAAGCGGATGGCTGCGAATCGGAGCGCGGATACCGTAAAGCACAGGAG 3956
4218 CGGTACGCCATTCGCGCGCAGCTCTTCAAGAAATACGCGGTAGCCAACTGTATGCC 4277
Db CGGTACGCCATTCGCGCGCAGCTCTTCAAGAAATACGCGGTAGCCAACTGTATGCC 3896
4278 TGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAAATCCAGAAAAAGCGCAATT 4337
QY

Db TGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAAAAAGCGCCATT 3836
QY TCCACCATGATATTCGCGAAGCAGGATCGCCATGGTCAAGCAGAGATCTCTCGCGTCG 4397
Db TCCACCATGATATTCGCGAAGCAGGATCGCCATGGTCAAGCAGAGATCTCTCGCGTCG 3776
QY GGATCGCGCTTGAAGCTTGGCGAAAGTTCGGCTGGCGAGCCCTGATGCTCTTCG 4457
Db GGATCGCGCTTGAAGCTTGGCGAAAGTTCGGCTGGCGAGCCCTGATGCTCTTCG 3716
QY TCCAGATCATCTGATCGACAGCGCTTCCATCCGAGTACGCTCGCTCGATGCGA 4517
Db TCCAGATCATCTGATCGACAGCGCTTCCATCCGAGTACGCTCGCTCGATGCGA 3656
QY TGTTCGCTTGGTGGTTCGAATGGCAGAGTAGCCGATCAAGCGTATGACGCCGCCATT 4577
Db TGTTCGCTTGGTGGTTCGAATGGCAGAGTAGCCGATCAAGCGTATGACGCCGCCATT 3596
QY GCATCAGCCATGATGATATCTTCTCGCAGAGCAAGTGAATGACAGGAGATCTTCG 4637
Db GCATCAGCCATGATGATATCTTCTCGCAGAGCAAGTGAATGACAGGAGATCTTCG 3536
QY CCGGCACTTCGCGCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAACTGACAGACA 4697
Db CCGGCACTTCGCGCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAACTGACAGACA 3476
QY GTTGGCAAGAAAGCGCGTCTGTCGACGACGATAGCGCGCTGCTCTCTCGAT 4757
Db GTTGGCAAGAAAGCGCGTCTGTCGACGACGATAGCGCGCTGCTCTCTCGAT 3416
QY TCATTGAGGGCACCGGACAGTCTTGAACAAAGAAACCGGGCGCCCTGCGCTGAC 4817
Db TCATTGAGGGCACCGGACAGTCTTGAACAAAGAAACCGGGCGCCCTGCGCTGAC 3356
QY AGCGGAAACACCGCGCATCAGAGCAGCGATGCTGTGTGTGCCAGTCATAGCCGAAT 3296
Db AGCGGAAACACCGCGCATCAGAGCAGCGATGCTGTGTGTGCCAGTCATAGCCGAAT 4937
4878 AGCCTCTCCACCAAGCGCGGAGAACTGCTGCAATCCCATCTTGTTCATCATGCGA 3236
3295 AGCCTCTCCACCAAGCGCGGAGAACTGCTGCAATCCCATCTTGTTCATCATGCGA 4937
QY AACGATCTCTCATCTCTCTCTGATCAGATCTTGTATCCCTGCGCCATCAGATCTTGGC 3235
Db AACGATCTCTCATCTCTCTGATCAGATCTTGTATCCCTGCGCCATCAGATCTTGGC 4997
QY GGCAAGAAAGCCATCCAGTTTACTTTGCAAGGCTTCCCAACCTTACAGAGGCGGCCA 5057
Db GGCAAGAAAGCCATCCAGTTTACTTTGCAAGGCTTCCCAACCTTACAGAGGCGGCCA 3116
QY GCTGGCAATTCGGTTTGGCTTGTGCTGCAATAAAACCGCCAGTCTAGCTATCGCCATGTA 5117
Db GCTGGCAATTCGGTTTGGCTTGTGCTGCAATAAAACCGCCAGTCTAGCTATCGCCATGTA 3056
QY AGCCCACTGCAAGCTACCTGCTTCTTGTGGCTTGTGCTTCTTGTCCAGATAGCC 5177
Db AGCCCACTGCAAGCTACCTGCTTCTTGTGGCTTGTGCTTCTTGTCCAGATAGCC 2996
QY CAGTGTGATCATTCATCCGGGTGACACCGTTTCTCGGAGTGGCTTCTTACGTTTC 5237
Db CAGTGTGATCATTCATCCGGGTGACACCGTTTCTCGGAGTGGCTTCTTACGTTTC 2936
QY CGCTTCCTTTAGCAGCCCTTGGCCCTGAGTGTCTCGGAGCGTG 5283
Db CGCTTCCTTTAGCAGCCCTTGGCCCTGAGTGTCTCGGAGCGTG 2890

RESULT 12
AY219861/c
LOCUS
DEFINITION Shuttle vector pEC-K19MECA2, complete sequence.
ACCESSION AY219861


```
VERSION      AY219861.1  GI:29164983
KEYWORDS
SOURCE       Shuttle vector pEC-K19MECA2
ORGANISM     Shuttle vector pEC-K19MECA2
REFERENCE    1 (bases 1 to 5869)
              artificial sequences; vectors.
AUTHORS      Kirchner,O. and Tauch,A.
TITLE        Tools for genetic engineering in the amino acid-producing bacterium
              Corynebacterium glutamicum
JOURNAL      Unpublished
REFERENCE    2. (bases 1 to 5869)
              Kirchner,O. and Tauch,A.
AUTHORS      Direct Submission
TITLE        Submitted (15-JAN-2003) Department of Genetics, University of
              Bielefeld, Universitaetsstrasse 25, Bielefeld D-33615, Germany
JOURNAL
FEATURES     source
              1..5869
                location/Qualifiers
                /organism="Shuttle vector pEC-K19MECA2"
                /mol_type="genomic DNA"
                /db_xref="taxon:224630"
                /note="Escherichia coli-Corynebacterium glutamicum shuttle
                vector"
              574..2037
                /gene="repA"
              574..2037
                /gene="repA"
                /note="replication protein"
                /codon_start=1
                /transl_table=11
                /product="repA"
                /protein_id="AA065228.1"
                /db_xref="GI:29164984"
                /translations="MTLADPQDTVTASAMKLSADLFDTHPEAMRCGSGRWTAEDRREL
                LAHLRGESFGSKTRPFSANWIKNPDTGETPKLYRAGSKALTRCOYVALTHAQAAY
                IVLDDVPSHQAGCKIEHVNQPVVALEKWARLEKAPANTGUNPLSKQLWLIDPV
                YAAAGKTSNNWLLAATTEMTYVFGADQAFSHLSRWPLHVSDDPTAYKNWQCHDRV
                DRALDMEIARTWGSQPKYIEQDFSSGRARIEAQAATAEAKALAILDASLVSAL
                DASGLDLDGVRVLTWPERARDETAFRHALTVGYQLKAAGERLKDAKIIDAYEVANV
                AOVAGDGRPEPLPAMRDRLTMARVGVYAKQPVVPAFRRVETOSRGRKALATWGR
                RGAATSNARWADPESYQAQETQRILAEANKRREMTGELLELRVKTAILDARSQSVAD
                PSTRELAGELVGSERRIQVKRKGALMEAKRGRPAEN"
                2680..2895
                /gene="per"
                /note="positive effector of replication"
                /codon_start=1
                /transl_table=11
                /product="per"
                /protein_id="AA065229.1"
                /db_xref="GI:29164985"
                /translations="MOLSDVALESDALDAVADLKTIVGFPRALDITTDAPASRDWASAA
                SDLETLVADLEELADELRARQRBQDAQ"
                3414..4208
                /gene="aph(3')-IIa"
                /note="aminoglycoside phosphotransferase"
                /codon_start=1
                /transl_table=11
                /product="Aph(3')-IIa"
                /protein_id="AA065230.1"
                /db_xref="GI:29164986"
                /translations="MLHQDGLHAGSPAANWVERLFGYDMAQQTIGCSDAAVFLSAGQR
                FVLVFKDLSGALNEQDEARLSWLAATGVPQCAVLVDVYVEAGRMEALLGVPGQDL
                LSHLPAEYKSIIMADARLHLTIDPATCFDFQAKHRIERARTMEAGLLVQDDIDLE
                EHQGLAPAEILFARKAMPDGEDLVTHGDACLPNTVNGRPSGFIDCGRLGVADRY
                QDIALATRDIAEELGGEWADRFVLVGLIAPDQSRIAFVRLIDFF"
                4593..4691
                /note="RP4-mob region"
              5454..5456
                /note="oriV"
              1325 a 1669 c 1683 g 1192 t

gene
CDS
misc_feature
rep_origin
BASE COUNT  1325 a 1669 c 1683 g 1192 t
```

ORIGIN

```
Query Match      39.5%; Score 2088.4; DB 12; Length 5869;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 141; Indels 231; Gaps 3;

QY 2709 GACGTAATCATGTCATAGCTGTTCCGTGCTGGAATTTGTTATCCGCTCACAATTCACCA 2768
DB 5869 GCGTAATCATGTCATAGCTGTTCCGTGCTGGAATTTGTTATCCGCTCACAATTCACCA 5810

QY 2769 CAACATACGAGCCGGAAGCATATAAGTGTAAAGCCCTGGGGTGCCTTAATGAGTGAGTAACT 2828
DB 5809 CAACATACGAGCCGGAAGCATATAAGTGTAAAGCCCTGGGGTGCCTTAATGAGTGAGTAACT 5750

QY 2829 CACATTAATTCGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGGAACCTGTGTCGCCAGCT 2888
DB 5749 CACATTAATTCGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGGAACCTGTGTCGCCAGCT 5690

QY 2889 GCATTAATGAATTCGGCCCAACCGCGGGGAGAGGGGGTTTGGGTATTTGGGCGCTCTTCGCG 2948
DB 5689 GCATTAATGAATTCGGCCCAACCGCGGGGAGAGGGGGTTTGGGTATTTGGGCGCTCTTCGCG 5630

QY 2949 TTCCTCGCTCACTGACTCGCTGCGTTCGGTTCGGTTCGGTTCGGGAGCGGTATCAGTCA 3008
DB 5629 TTCCTCGCTCACTGACTCGCTGCGTTCGGTTCGGTTCGGGAGCGGTATCAGTCA 5570

QY 3009 CTCAAAGCGGTAATACCGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGAACATGTG 3068
DB 5569 CTCAAAGCGGTAATACCGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGAACATGTG 5510

QY 3069 AGCAAAAGCCAGCAAAAGGCGCAGAAACCGTAAGAACCCCGTGTGTGGCGTTTTCCTCA 3128
DB 5509 AGCAAAAGCCAGCAAAAGGCGCAGAAACCGTAAGAACCCCGTGTGTGGCGTTTTCCTCA 5450

QY 3129 TAGGTCCTCGCCCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGGAA 3188
DB 5449 TAGGTCCTCGCCCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGGAA 5390

QY 3189 CCCGACAGAGCTATAAAGATACCAAGCGGCTTCCCGCTCGGAAGCTCCCTCGTGGCGCTCC 3248
DB 5389 CCCGACAGAGCTATAAAGATACCAAGCGGCTTCCCGCTCGGAAGCTCCCTCGTGGCGCTCC 5330

QY 3249 TGTTCCGACCTTGGCGCTTACGGATACCTGTCGCGCTTTCCTCGTCCGGAAGCGTGGC 3308
DB 5329 TGTTCCGACCTTGGCGCTTACGGATACCTGTCGCGCTTTCCTCGTCCGGAAGCGTGGC 5270

QY 3309 GCTTTCTCATAGCTCAGCGCTAGTGTATCTCAGTTCGGTGTAGTGTGCTTCCGCTCAAGCT 3368
DB 5269 GCTTTCTCATAGCTCAGCGCTAGTGTATCTCAGTTCGGTGTAGTGTGCTTCCGCTCAAGCT 5210

QY 3369 GGGCTGTGTGACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCCGTAACTATCG 3428
DB 5209 GGGCTGTGTGACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCCGTAACTATCG 5150

QY 3429 TCTTGAGTCCAAACCGGTAAGACACGACTTATCCCACTGCGCAGCACCCTGCTTACAG 3488
DB 5149 TCTTGAGTCCAAACCGGTAAGACACGACTTATCCCACTGCGCAGCACCCTGCTTACAG 5090

QY 3489 GATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCCTTGAAGTGGTGGCTTAACCTA 3548
DB 5089 GATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCCTTGAAGTGGTGGCTTAACCTA 5030

QY 3549 CGGCTACACTAGAGACAGTATTGTTGTTCTGCGCTCTGCTGAAGCCAGTACTCTTCGG 3608
DB 5029 CGGCTACACTAGAGACAGTATTGTTGTTCTGCGCTCTGCTGAAGCCAGTACTCTTCGG 4970

QY 3609 AAAAAGAGTTCGTAGCTCTTGATCCGCAAAACCAACCCGCTGCTGAGCGGTGGTGTCTT 3668
DB 4969 AAAAAGAGTTCGTAGCTCTTGATCCGCAAAACCAACCCGCTGCTGAGCGGTGGTGTCTT 4910

QY 3669 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAAAGATCTTTGATCTT 3728
DB 4909 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAAAGATCTTTGATCTT 4850
```



```

source
1. .5695
/organism="Cloning vector pEC-K18mob2"
/mol_type="genomic DNA"
/db_xref="taxon:179875"
/lab_host="Corynebacterium glutamicum"
184_..1647
/gene="rep"
184_..1647
/gene="rep"
/notes="Rep"
/codon_start=1
/transl_table=11
/product="replication protein"
/protein_id="AAL38580.1"
/db_xref="GI:17386068"
/translation="MTLADPQDVTASAKLSADLFDTHPEMRCGSGRWTAEDRREL
LAHLGRESQGSKTRDFASAKINKNPTDGTQPKLYRAGSKALTRCOYVALTHQAHA
VLVIDIPVSGAGKIEHNVQYVAILKWARLEKAPAWIGVNLPSGKQLWLIDPV
YAAAGTSPNRLLAATTEMTVRFGADQAFSHLSRWPLHVSDDPTAYKWHQCHDVR
DLADIMEATRTWTGSKPKKYTEQDFSSGRARI EAAORATAEKAKALIIDLASLP
DASGDLIDGVRVLVTPERAPDFTA PHALTVGYQLKAAGERLAKDAKIIDA YEVAVV
AQAAGDREPDLPANRDLRUTMARRVRGVYAKQPVVPARRVETQSSRGRKALATWGR
RGATSNARMAWDPESKYQAEHQRLAEANKRHEMTGELLELELVKVTAILDARSQS
PSTRELAGLIGWSERRIQVQRKALGMEAKRGPRPRAEN"
2290..2505
/gene="per"
2290..2505
/gene="per"
/notes="Per"
/codon_start=1
/transl_table=11
/product="positive effector of replication"
/protein_id="AAL38581.1"
/db_xref="GI:17386069"
/translation="MOLSDVLESDDLDAADVLDKTVIGFPRALDITDAPASRDWASA
SDELTIVADLEELADELARAKQROEDAQ"
3024..3818
/gene="aph(3')-IIa"
3024..3818
/gene="aph(3')-IIa"
/notes="Aph(3')-IIa"
/codon_start=1
/transl_table=11
/product="kanamycin resistance protein"
/protein_id="AAL38582.1"
/db_xref="GI:17386070"
/translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVRLSAQGR
PVLVFKTDLSGALNEIQDEARLSWLATGVCVAALVDVYTEAGRWLLIGVEGDDLL
LSHLAPAEKVSIMADNRHLHIDPATCFDQAKHRIERATREAGHLVDDQDDLE
EHQGLAPAEFALRKAMRPGEDLVVTHGDACLPNINVENRFSFGIDCGRLGVADRY
QDALATRDIAEELGGSWADRFVLVYGINAPDSPQRIATFYELLDEFF"
4203..4301
/misc_feature
/note="RP4 mob region"
5064..5066
/rep_origin
/note="oriV of Escherichia coli"
BASE COUNT 1285 a 1624 c 1638 g 1148 t
ORIGIN

Query Match 39.5%; Score 2088; DB 12; Length 5695;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 2433; Conservative 0; Mismatches 140; Indels 231; Gaps 3;

Qy 2711 CGTAATCATGCTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 2770
Db 5477 CGTAATCATGCTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 5418

Qy 2771 ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTGAGCTAACTCA 2830
Db 5417 ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTGAGCTAACTCA 5358

Qy 2831 CATTAATTGCGTTGCGTCACTGCCCGCTTTTCCAGTCGGGAAAACCTGTCGTCGCGAGCTGC 2890
Db 5357 CATTAATTGCGTTGCGTCACTGCCCGCTTTTCCAGTCGGGAAAACCTGTCGTCGCGAGCTGC 5298

```


misc_feature		1763..1861	/note="BP4-mob region"			
rep_origin		2624..2626	/note="oriV"			
BASE COUNT		665 a	841 c	874 g	655 t	
ORIGIN						
Qy	2721	GTCTAGCTGTTTCCGTGTGAAATTTGTTATCCGCTCACAATTCACACACATACGAGC	2780			
Db	3027	GTCTAGCTGTTTCCGTGTGAAATTTGTTATCCGCTCACAATTCACACACATACGAGC	2968			
Qy	2781	CGGAACATAAAGTGAAGCCCTGGGGTGCCTAAATGAGTGAGCTAACTCACATTAATTCG	2840			
Db	2967	CGGAACATAAAGTGAAGCCCTGGGGTGCCTAAATGAGTGAGCTAACTCACATTAATTCG	2908			
Qy	2841	GTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTGTCGACGCTGCAATTAATGAAT	2900			
Db	2907	GTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTGTCGACGCTGCAATTAATGAAT	2848			
Qy	2901	CGGCCAACGGCGGGGAGAGGGGTTTGGTATTTGGCGCTCTTCGGCTTCTCGCTCAC	2960			
Db	2847	CGGCCAACGGCGGGGAGAGGGGTTTGGTATTTGGCGCTCTTCGGCTTCTCGCTCAC	2788			
Qy	2961	TGACTCGCTCGCTCGCTGTTTCGGCTCGGGGAGCGGTATCAGCTCACTCAAAGCGCGT	3020			
Db	2787	TGACTCGCTCGCTCGCTGTTTCGGCTCGGGGAGCGGTATCAGCTCACTCAAAGCGCGT	2728			
Qy	3021	AATACGGTTATCACAGATCAGGGATTAACGAGGAAGACATGTGAGAAAGGCCA	3080			
Db	2727	AATACGGTTATCACAGATCAGGGATTAACGAGGAAGACATGTGAGAAAGGCCA	2668			
Qy	3081	GCAAAAGGCCAGAACCGTAAAAAGCCGGTGTGCTGGCGTTTTTCCATAGGCTCCGCC	3140			
Db	2667	GCAAAAGGCCAGAACCGTAAAAAGCCGGTGTGCTGGCGTTTTTCCATAGGCTCCGCC	2608			
Qy	3141	CCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACT	3200			
Db	2607	CCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACT	2548			
Qy	3201	ATAAAGATACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCT	3260			
Db	2547	ATAAAGATACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCT	2488			
Qy	3261	GCCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAG	3320			
Db	2487	GCCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAG	2428			
Qy	3321	CTCAGCTGTAGTATCTCAGTTTCGGGTAGGTGCTTCGCTCAAGCTGGGCTGTGCA	3380			
Db	2427	CTCAGCTGTAGTATCTCAGTTTCGGGTAGGTGCTTCGCTCAAGCTGGGCTGTGCA	2368			
Qy	3381	CGAACCCCGCTTCAGCCCGACCGCTCGGCTTATCCGGTAACTATCGTTCGAGTCCAA	3440			
Db	2367	CGAACCCCGCTTCAGCCCGACCGCTCGGCTTATCCGGTAACTATCGTTCGAGTCCAA	2308			
Qy	3441	CCCGGTAAAGACACGACTTATCGCCACTGGCAGCGACCTGGTAAACGAGTATGACGAGC	3500			
Db	2307	CCCGGTAAAGACACGACTTATCGCCACTGGCAGCGACCTGGTAAACGAGTATGACGAGC	2248			
Qy	3501	GAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGCTACACTAG	3560			
Db	2247	GAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGCTACACTAG	2188			
Qy	3561	AAGAACAGTATTTGGTATCTGGCTCTGCTGAGCCAGTTACCTTCGAAAAAGATTGG	3620			
Db	2187	AAGAACAGTATTTGGTATCTGGCTCTGCTGAGCCAGTTACCTTCGAAAAAGATTGG	2128			
Qy	3621	TAGCTCTTGATCCGGCAACAAACACCGCTGTGGGTGGTTTTTCTGTTTTCGAAGCA	3680			
Db						
Db	2127	TAGCTCTTGATCCGGCAACAAACACCGCTGTGGGTGGTTTTTCTGTTTTCGAAGCA	2068			
Qy	3681	GCAGTTTACCGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCTTTCTACGGGTC	3740			
Db	2067	GCAGTTTACCGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCTTTCTACGGGTC	2008			
Qy	3741	TGACGCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTTCATGAGATTATCGTGC	3798			
Db	2007	TGACGCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTTCATGAGATTATCGTGC	1948			
Qy	3799	-----	3798			
Db	1947	TAACTCTGCTTCGGGTTCATTTATAGCAATTTTTCGGTATATCCATCTTTTTCGACGA	1888			
Qy	3799	-----	3798			
Db	1887	TATACAGGATTTTGCCAAAGGTTTCGTGTAGACTTTTCTCTTGGTGTATCCAAACGGCGTCAG	1828			
Qy	3799	-----ACCAAGCGGCCATCGTCCCTCCCACTCTGCAAGTTTCGGGGCATGGATCGC	3852			
Db	1827	CGGGCAGGATAGTGAAGTAGGCCCAACCCGCGAGCGGGTGTCTTCTTCTCACTGTCCT	1768			
Qy	3853	GGATAGCCGCTGCTGTGTTCCTCGATGCCGACGGAATTTGCATGCGCGGTAGAACTCGGC	3911			
Db	1767	TATTCGACCTGGCGGTCTCAACGGGAATCTCTCTGCGAGCTGGCGGTACCGCC	1708			
Qy	3912	-----GAGTCTGCTCCAGCTCAGGCCTCAGGCAGCTGAACCACTCCGAGGGGATC	3959			
Db	1707	GGCGTAAACAGATGAGGCAAGCGGATGGTGTAAACCAAGCCCAACGAGAGGGGAGC	1648			
Qy	3960	GAGCCCGGGTGGCGAAGAACTCCAGCTGAGATCCCGCTGGAGGATCATCCAGC	4018			
Db	1647	CCACTATCAAGGTGTACTGCTTCAGACGAAACGAGAGCGATTGAGGAAAGCGCGC	1588			
Qy	4019	-----	4018			
Db	1587	CGCGCCGCGCATAGCCTGTCCGCTACCTGTCCGCTCGGCCAGGCTCAAAATCAGC	1528			
Qy	4019	-----CGCGCTCCCGGAAAAACGATTCCGAAAGCCCAA	4049			
Db	1527	GGCGTCTGTGACTATAGCAGCTCCGCGAGGGCGTCCCGGAAACGATTCCGAAAGCCCAA	1468			
Qy	4050	CCTTTCATAGAAAGCGCGGTGGAATCGAAATCTCTGTGATGGAGGTGGCGTGGCTTG	4109			
Db	1467	CCTTTCATAGAAAGCGCGGTGGAATCGAAATCTCTGTGATGGAGGTGGCGTGGCTTG	1408			
Qy	4110	GTGCTCATTTTCGAACCCAGCTCCGCTCAGAGAACTCGTCAAGAGCGGATAGAAG	4169			
Db	1407	GTGCTCATTTTCGAACCCAGCTCCGCTCAGAGAACTCGTCAAGAGCGGATAGAAG	1348			
Qy	4170	GCGATGCGCTCGGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTACGCCAT	4229			
Db	1347	GCGATGCGCTCGGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTACGCCAT	1288			
Qy	4230	TCGCGCCCAAGCTCTTTCAGCAATATCAGGGTAGCCACGCTATGCTCTGATACGGTCC	4289			
Db	1287	TCGCGCCCAAGCTCTTTCAGCAATATCAGGGTAGCCACGCTATGCTCTGATACGGTCC	1228			
Qy	4290	GCACACCCAGCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATA	4349			
Db	1227	GCACACCCAGCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATA	1168			
Qy	4350	TTCCGCAAGAGCGGATCCCATATGGGTTCAGAGATCTTCGCGCTGGGATCGCGCC	4409			
Db	1167	TTCCGCAAGAGCGGATCCCATATGGGTTCAGAGATCTTCGCGCTGGGATCGCGCC	1108			
Qy	4410	TTGAGCCTTGGGAAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTCCAGATCATCC	4469			
Db	1107	TTGAGCCTTGGGAAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTCCAGATCATCC	1048			
Qy	4470	TGATCGAACAAGACCGGCTTCCATCCAGTACGTAAGTCTGCTCGATGCGGATGTTTGGTTGG	4529			
Db	1047	TGATCGAACAAGACCGGCTTCCATCCAGTACGTAAGTCTGCTCGATGCGGATGTTTGGTTGG	988			

QY	3261	GCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGAAAGCTGGCGCTTCTCATAG	3320
DB	2505	GCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGAAAGCTGGCGCTTCTCAATG	2446
QY	3321	CTCACGCTGAGGTACTCAGTTCGCTGAGTTCGTTCTCCCTCCAGCTCGGCTGTGTGCA	3380
DB	2445	CTCACGCTGAGGTACTCAGTTCGCTGAGTTCGTTCTCCCTCCAGCTCGGCTGTGTGCA	2386
QY	3381	CGAACCCCGCTTCAGCCCGACCGCTTCGCTTATCCGGTAACTATCGCTTTGAGTCCAA	3440
DB	2385	CGAACCCCGCTTCAGCCCGACCGCTTCGCTTATCCGGTAACTATCGCTTTGAGTCCAA	2326
QY	3441	CCGGTTAAGACACGACTTATCGCACTGGCAGCAGCACCTGGTAAACAGATTACGACGC	3500
DB	2325	CCGGTTAAGACACGACTTATCGCACTGGCAGCAGCACCTGGTAAACAGATTACGACGC	2266
QY	3501	GAGGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTTAATACTACGGCTACACTAG	3560
DB	2265	GAGGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTTAATACTACGGCTACACTAG	2206
QY	3561	AAGAAACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTCCTTCGGAAAGAGTTCG	3620
DB	2205	AAGAAACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTCCTTCGGAAAGAGTTCG	2146
QY	3621	TAGCTCTTGATCCGGCAAAACAAACCCGCTGCTAGCGGTGTTTTTTTGTTCGAAGCA	3680
DB	2145	TAGCTCTTGATCCGGCAAAACAAACCCGCTGCTGCTAGCGGTGTTTTTTGTTCGAAGCA	2086
QY	3681	GCAGATTACGCGCAGAAAAGAGATCTCAAGAGATCCTTTGATCTTTTCTACGGGTC	3740
DB	2085	GCAGATTACGCGCAGAAAAGAGATCTCAAGAGATCCTTTGATCTTTTCTACGGGTC	2026
QY	3741	TGACGCTCAGTGGAAACGAAACTCAGGTTAAGGGATTTTGGTCAATGAGATTATCGTCG	3798
DB	2025	TGACGCTCAGTGGAAACGAAACTCAGGTTAAGGGATTTTGGTCAATGAGATTATCGTCG	1966
QY	3799	-----	3798
DB	1965	TAACCTGCTTCGGGTCATATPAGCGATTTTTCGGTATATCCATCCTTTTTCGCACGA	1906
QY	3799	-----	3798
DB	1905	TATACAGGATTTTGCACAAAGGTTTCGTGTAGACTTTCTTGGTGATATCCAAACGGCTCAG	1846
QY	3799	-----ACAAAGGGCCATCGTCCCTCCCACTCTGTCAGTTTCGGGGCATCGATCGCG	3852
DB	1845	CGGGGCAGGATAGGTGAAGTAGGAGGCCACCCGCGAGCGGGTGTTCTTCTCATGTGCCCT	1786
QY	3853	GGATAGCGCTGCTGGTTTCTCGATGCCACGGAATTTGCACCTCCCGGTAGAACTCCGC	3911
DB	1785	TATTTCGACCTGCGGTGCTCAACGGGAATCCTGCTCTCGAGGCTGGCGGCTACCGCC	1726
QY	3912	-----GAGGTGCTCGAGCTCAGCAGCAGCTGAAACCAACTCCGAGGGGATC	3959
DB	1725	GGCGTAAACAGTAGGGCAACGCGATGGCTGATGAAACCAAGGCCAACAGGAAGGCGCAGC	1666
QY	3960	GAGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGC	4018
DB	1665	CCACCTATCAAGGTGTACTGCTTCCAGACGAAAGAGCGGATTGAGGAAAGGCGCG	1606
QY	4019	-----	4018
DB	1605	GCGCCCGGCATGAGCCTGTCCGCTACCTGCTGCGCGCTCGGCCAGGCTACAAAATCAGC	1546
QY	4019	-----CGCGCTCCCGGAAACAGATTCGGAAGCCCAA	4049
DB	1545	GGCGTGTGGACTATGACGACGTCGCGAGGGCTCCCGGAAACAGATTCCGAAGCCCAA	1486
QY	4050	CTTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTGATGGCAGGTTGGCGTCGCTTG	4109
DB	1485	CTTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTGATGGCAGGTTGGCGTCGCTTG	1426

QY	4110	GTCCGCTCATTTCCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGGGATAGAAG	4166
DB	1425	GTCCGCTCATTTCCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGGGATAGAAG	1366
QY	4170	GGATGCGCTGCGAATCCGGAGCGGGATACCGTAAAGCACAGGAAGCGGTCAGCCCAT	4229
DB	1365	GGATGCGCTGCGAATCCGGAGCGGGATACCGTAAAGCACAGGAAGCGGTCAGCCCAT	1306
QY	4230	TGCGCGCCAAGCTCTTTCAGCAATATCATCGGGTAGCCACGCTATGTCCTGTATAGCGGTCC	4289
DB	1305	TGCGCGCCAAGCTCTTTCAGCAATATCACGGGTAGCCACGCTATGTCCTGTATAGCGGTCC	1246
QY	4290	GCCACACCCAGCGGCCACAGTCGATGAATCCAGAAAAAGCGCCATTTTCCACCATGATA	4349
DB	1245	GCCACACCCAGCGGCCACAGTCGATGAATCCAGAAAAAGCGCCATTTTCCACCATGATA	1186
QY	4350	TTTGGCAAGACGACATCGCATGGGTACGACGAGATCTCGCGCTGGGGCATGCGCGCC	4409
DB	1185	TTTGGCAAGACGACATCGCATGGGTACGACGAGATCTCGCGCTGGGGCATGCGCGCC	1126
QY	4410	TTGAGCCTCGGCAACAGTTTCGGCTGCGGAGCCCTGATGCTCTTCGTCAGATCATCC	4469
DB	1125	TTGAGCCTCGGCAACAGTTTCGGCTGCGGAGCCCTGATGCTCTTCGTCAGATCATCC	1066
QY	4470	TGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTTCGTTGG	4529
DB	1065	TGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTTCGTTGG	1006
QY	4530	TGCTCAATTGGGAGGTAGCCGGATCAAGCGTATCAGCGCGCGCATTCGATCAGGCATG	4589
DB	1005	TGCTCAATTGGGAGGTAGCCGGATCAAGCGTATCAGCGCGCGCATTCGATCAGGCATG	946
QY	4590	ATGGAATCTTTCGCGCAGGAGCAAGGTGAGATGACAGGAGATCTCGCCCGGCATCTCG	4649
DB	945	ATGGAATCTTTCGCGCAGGAGCAAGGTGAGATGACAGGAGATCTCGCCCGGCATCTCG	886
QY	4650	CCCAATAGCAGCCAGTCCCTTCCCGTTTCAAGTGACAACGTTCGAGCAGCGTGGCAAGGA	4709
DB	885	CCCAATAGCAGCCAGTCCCTTCCCGTTTCAAGTGACAACGTTCGAGCAGCGTGGCAAGGA	826
QY	4710	ACGCCGCTGTGCCACGACAGTACCGCGCTGCTCGTCTGTGAGTTTCAATTCAGGGCA	4769
DB	825	ACGCCGCTGTGCCACGACAGTACCGCGCTGCTCGTCTGTGAGTTTCAATTCAGGGCA	766
QY	4770	CCGGAACAGTTCGTCTTGACAAAAAAGAAACGGGCGCCCTTCGCTGACAGCGGAAACAG	4829
DB	765	CCGGAACAGTTCGTCTTGACAAAAAAGAAACGGGCGCCCTTCGCTGACAGCGGAAACAG	706
QY	4830	GGCGCATCAGACAGCCGATTTGCTGTGTGCCAGTTCATAGCCGAATAGCCTCTCCACC	4889
DB	705	GGCGCATCAGACAGCCGATTTGCTGTGTGCCAGTTCATAGCCGAATAGCCTCTCCACC	646
QY	4890	CAAGCGGCCGGAAACCTGCGTCCAATCCATCTTGTTCATATGCGAAACGATCTCTCAT	4949
DB	645	CAAGCGGCCGGAAACCTGCGTCCAATCCATCTTGTTCATATGCGAAACGATCTCTCAT	586
QY	4950	CCTGTCTCTTGATCAGATCTTTGATCCGCTGGCGCATCAGATCCTTGGCGCAAGAAAGCC	5009
DB	585	CCTGTCTCTTGATCAGATCTTTGATCCGCTGGCGCATCAGATCCTTGGCGCAAGAAAGCC	526
QY	5010	ATCCAGTTTACTTTGACGGGCTTCCCAACCTTACACAGAGGGCGCCCGAGCTGGCAATTC	5069
DB	525	ATCCAGTTTACTTTGACGGGCTTCCCAACCTTACACAGAGGGCGCCCGAGCTGGCAATTC	466
QY	5070	GGTTTCGCTTCTGTCATAAAAACGGCCAGTCTAGCTATCGCCATGTAGCCCCACTGCAA	5129
DB	465	GGTTTCGCTTCTGTCATAAAAACGGCCAGTCTAGCTATCGCCATGTAGCCCCACTGCAA	406
QY	5130	GCTACCTGCTTCTCTTTTGGCTTGCGTTTTTCCCTTGTCCAGATAGCCCACTAGCTGACA	5189
DB	405	GCTACCTGCTTCTCTTTTGGCTTGCGTTTTTCCCTTGTCCAGATAGCCCACTAGCTGACA	346
QY	5190	TTTCATCCGGGGTCAGCACCGTTTTCTCGGACTGGCTTTCTTACGTTTCCGCTTCCCTTAG	5249

Db 345 TTATACCGGGTTCAGACCGTTTCTGCGACTGGCTTTCTACGTGTTCCGCTTCCTTAG 286

Qy 5250 CAGCCCTTGGCCCTGAGTGTCTTGGCGAGCGTG 5283

Db 285 CAGCCCTTGGCCCTGAGTGTCTTGGCGAGCGTG 252

Search completed: February 15, 2004, 16:21:29
Job time : 18649 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 05:44:19 ; Search time 1245 Seconds
(without alignments)

11454.724 Million cell updates/sec

Title: US-09-921-143-36

Perfect score: 5283

Sequence: 1 aagttgacctatgcgact.....tgagtgttgcggcagcgtg 5283

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5283	100.0	5283	24	ABK10062
2	2240	42.4	8349	21	AA15565
3	2017.2	38.2	6136	19	AAV29673
C 4	1793.4	33.9	4800	20	AAV52020
C 5	1652.4	31.3	5594	19	AAV33629
C 6	1634.4	30.9	5285	22	AAH74865
C 7	1618.2	30.6	6561	19	AAV33630
8	1603.8	30.4	2192	24	ABQ78294

9	1603.8	30.4	3534	24	ABQ78293	Nucleotide sequenc
10	1603.8	30.4	3599	19	AAV50428	Plasmid pIG0552 lo
11	1603.8	30.4	3599	19	AAV40796	Actual sequence of
12	1602.2	30.3	3600	19	AAV50427	Plasmid pIG0552 up
13	1602.2	30.3	3600	19	AAV40795	Expected sequence
14	1602.2	30.3	3600	19	AAV40795	Plasmid pIG0335 DN
15	1600	30.3	3566	20	AAZ40417	Plasmid pIN0961 co
C 16	1571.6	29.7	4665	22	AAF30342	Construct BIOTOP-
C 17	1570.8	29.7	5068	20	AAH84027	MP9 promoter GFP
18	1570	29.7	4517	22	AAH24427	pBKMV phagemid.
19	1570	29.7	4518	19	AAV33626	Plasmid pBK-QWV DN
20	1570	29.7	4886	24	AAZ42469	Human aequorin-pBk
C 21	1570	29.7	5031	20	AAZ23355	EGFP/DRM fusion pr
C 22	1570	29.7	5054	20	AAZ23354	EGFP/DRM fusion pr
C 23	1570	29.7	5130	20	AAZ23353	EGFP/DRM fusion pr
C 24	1570	29.7	5168	20	AAZ23351	EGFP/DRM fusion pr
C 25	1570	29.7	5619	21	AAA07773	DNA sequence of pl
C 26	1570	29.7	5627	21	AAA07778	DNA sequence of pl
C 27	1570	29.7	5775	25	ABZ24716	Murine potassium c
C 28	1570	29.7	6121	21	AAA07779	DNA sequence of pl
C 29	1570	29.7	6139	21	AAZ40023	Interleukin-12 fus
C 30	1570	29.7	7655	21	AAA07847	Hs-UNC-53/1 fragm
C 31	1570	29.7	8140	21	AAA07845	Hs-UNC-53/3 fragm
C 32	1570	29.7	11842	21	AAA07846	Hs-UNC-53/3 DNA c
C 33	1569.8	29.7	4055	22	AAI68460	Plasmid pSV-HIII-C
C 34	1569.8	29.7	4853	22	AAZ10015	Plasmid pSFM 8 to
C 35	1569.8	29.7	4892	24	AAZ47954	Modified yeast str
C 36	1569.8	29.7	4892	25	ACC44640	Vector pBRedIN1 n
C 37	1569.8	29.7	4892	25	ACC44640	Vector pBRedIN1 n
C 38	1569.8	29.7	4700	21	AAA48561	Anemonia majano mu
C 39	1569.8	29.7	4700	21	AAA48561	Zoanthus sp. novel
C 40	1569.8	29.7	4727	25	ABQ80035	VSV M C-terminal/H
C 41	1569.8	29.7	4729	21	AAA07772	DNA sequence of ve
C 42	1569.8	29.7	4735	21	AAA07777	DNA sequence of ve
C 43	1569.8	29.7	4740	25	ABQ80036	Hifi-alpha ODD/VS
C 44	1569.8	29.7	4821	21	AAA50946	Anemonia majano mu
C 45	1569.8	29.7	4821	21	AAA50947	Anemonia majano mu

ALIGNMENTS

RESULT 1
ABK10062
ID ABK10062 standard; DNA; 5283 BP.
XX
XX
AC ABK10062;
XX
XX
DT 21-MAY-2002 (first entry)
XX
DE Expression vector construct pVGI.1 containing VEGF-2 insert.
XX
KW Human; vascular endothelial growth factor 2; VEGF-2; ds; pVGI.1;
KW chronic limb ischaemia; myocardial ischaemia; autoimmune disorder; pHE4a;
KW allergic reaction; organ rejection; inflammatory condition; arrhythmia;
KW hyperproliferative disorder; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cardiovascular disorder; embolism;
KW heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.
OS Synthetic.
XX
XX
FN WO200211769-A1.
XX
XX
PD 14-FEB-2002.
XX
XX
PF 03-AUG-2001; 2001WO-US24658.
XX
XX
PR 04-AUG-2000; 2000US-223276P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Coleman TA;
XX

DR WPI: 2002-217153/27.

XX Isolated nucleic acid having expression vector construct with vascular

PT endothelial growth factor-2 insert, useful for treating chronic limb

PT ischaemia or myocardial ischaemia, autoimmune disorders and allergic

PT conditions

XX

XX Disclosure; Fig 31; 241pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising pVGI.1

CC expression vector construct containing the vascular endothelial growth

CC factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host

CC cell by transducing, transforming or transfecting a host cell with the

CC DNA and for treating a patient having chronic limb ischaemia or

CC myocardial ischaemia, or a disease or disorder selected from autoimmune

CC disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or

CC conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.

CC Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),

CC diseases due to viral, bacterial, fungal or parasitic infection,

CC cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve

CC diseases, aneurysms, arterial occlusive disorders and embolism. This

CC sequence represents the pVGI.1 expression vector containing the VEGF-2

CC insert.

XX

SQ Sequence 5283 BP; 1244 A; 1422 C; 1417 G; 1200 T; 0 other;

Query Match 100.0%; Score 5283; DB 24; Length 5283;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTTGACCTTATCGCACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT 60

DB 1 AAGCTTGACCTTATCGCACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT 60

QY 61 TACCATGGTGATCGCGTTTGGCAGTACATCAATGGCGGTGGATACGGTTTGACTCAG 120

DB 61 TACCATGGTGATCGCGTTTGGCAGTACATCAATGGCGGTGGATACGGTTTGACTCAG 120

QY 121 GGGATTTCGAAGTCTCCACCCCACTGACGTCAATGGGAGTTTGTTTGGCACCAAAATCA 180

DB 121 GGGATTTCGAAGTCTCCACCCCACTGACGTCAATGGGAGTTTGTTTGGCACCAAAATCA 180

QY 181 ACGAGACTTTCCAAATGTCGTAAACAATCCGCCCACTGACGCAATGGCGGTAGGCG 240

DB 181 ACGAGACTTTCCAAATGTCGTAAACAATCCGCCCACTGACGCAATGGCGGTAGGCG 240

QY 241 AACATGCTTATGTAAACGGTGAAGTTAGCAACATGCTTTAAGGAGAGAAAAAGCA 306

DB 241 AACATGCTTATGTAAACGGTGAAGTTAGCAACATGCTTTAAGGAGAGAAAAAGCA 306

QY 301 CATGCCGATTTGGTGGGAGTAAAGTGGTATGATCGTGTATGATCGTCCCTGTTTAGGAG 360

DB 301 CATGCCGATTTGGTGGGAGTAAAGTGGTATGATCGTGTATGATCGTCCCTGTTTAGGAG 360

QY 361 GCAACAGACGGGTCTAAACAGGATTTGACGAAACCACTGAATTCGCAATTGACAGATATT 420

DB 361 GCAACAGACGGGTCTAAACAGGATTTGACGAAACCACTGAATTCGCAATTGACAGATATT 420

QY 421 GTATTTAAGTGCCCGACTCGATACAAATAAAGCCATTGACATTACCACTTGGTGTG 480

DB 421 GTATTTAAGTGCCCGACTCGATACAAATAAAGCCATTGACATTACCACTTGGTGTG 480

QY 481 CACCTGGTTGGATCGATCCATCATGCACTCGTGGGCTTCTTCTCTGTGGCGGTGTTCT 540

DB 481 CACCTGGTTGGATCGATCCATCATGCACTCGTGGGCTTCTTCTCTGTGGCGGTGTTCT 540

QY 541 CTGCTGCGCGCTGCGTGTCTCCGGGTCTCTCGCAGGCGCCGCGCGCGCGCGCGCTTC 600

DB 541 CTGCTGCGCGCTGCGTGTCTCCGGGTCTCTCGCAGGCGCCGCGCGCGCGCGCGCTTC 600

QY 601 GAGTCCGGAATCGAATCTCTCGAAGCGGAGCCGACGCGGTGAGGCAACGGCTTATGCA 660

DB 601 GAGTCCGGAATCGAATCTCTCGAAGCGGAGCCGACGCGGTGAGGCAACGGCTTATGCA 660

Db	3901	TAGAACTCCGCGAGGTGCTCCAGCCTCAGCGAGAGACTGAAACCAACTCCGAGGGGATCG	3960
Qy	3961	AGCCCGGGTGGCGCAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGC	4020
Db	3961	AGCCCGGGTGGCGCAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGC	4020
Qy	4021	GCCTCCGGGAAAACGATTCCGAAGCCCAACCTTTATAGAAAGCGCGGTGGAATCGAAA	4080
Db	4021	GCCTCCGGGAAAACGATTCCGAAGCCCAACCTTTATAGAAAGCGCGGTGGAATCGAAA	4080
Qy	4081	TCTCGTGATCGCAGGTGTGGCGCTCGCTTGCTCGGTCAATTTGCAACCCCAAGATCCCGCTC	4140
Db	4081	TCTCGTGATCGCAGGTGTGGCGCTCGCTTGCTCGGTCAATTTGCAACCCCAAGATCCCGCTC	4140
Qy	4141	AGAAGAACTCGTCAAGAAGCGGATAGAAAGCGATGCGCTCGGAATCGGAGCGCGGATAC	4200
Db	4141	AGAAGAACTCGTCAAGAAGCGGATAGAAAGCGATGCGCTCGGAATCGGAGCGCGGATAC	4200
Qy	4201	CGTAAAGCAGAGGAAGCGGTAGCCATTTCGCGCGCCCAAGCTCTTTCAGCAATATCACGGG	4260
Db	4201	CGTAAAGCAGAGGAAGCGGTAGCCATTTCGCGCGCCCAAGCTCTTTCAGCAATATCACGGG	4260
Qy	4261	TAGCCAAACGCTATGTCTCTGATAGCGGTTCGCCACAACCCAGCCGGCCACAGTCGATGAATC	4320
Db	4261	TAGCCAAACGCTATGTCTCTGATAGCGGTTCGCCACAACCCAGCCGGCCACAGTCGATGAATC	4320
Qy	4321	CAGAAAAGCGGCCATTTTCCACCATGATATTCCGCGAGCAGGCGATCGCCATGGTCAACGA	4380
Db	4321	CAGAAAAGCGGCCATTTTCCACCATGATATTCCGCGAGCAGGCGATCGCCATGGTCAACGA	4380
Qy	4381	CGAGATCCTCGCCGTCGGCGATGCGCGCTTGAGCCTTGGCGAACAAGTTCCGCTGGCGCGA	4440
Db	4381	CGAGATCCTCGCCGTCGGCGATGCGCGCTTGAGCCTTGGCGAACAAGTTCCGCTGGCGCGA	4440
Qy	4441	GCCCTCGATGCTTCTGTCAGATCATCTGTATGCAACAAGCCGGCTTCATCCGAGTAC	4500
Db	4441	GCCCTCGATGCTTCTGTCAGATCATCTGTATGCAACAAGCCGGCTTCATCCGAGTAC	4500
Qy	4501	GTGCTCGCTCGATCGCATGTTTTCCGTTGGTGGTTCGAATGGCGAGGTAGCCGGATCAAGCG	4560
Db	4501	GTGCTCGCTCGATCGCATGTTTTCCGTTGGTGGTTCGAATGGCGAGGTAGCCGGATCAAGCG	4560
Qy	4561	TATCAGCGCCGCGCATTCGATCAGCCATGATGATATCTTTCTCGCAGGAGCAAGGTGAG	4620
Db	4561	TATCAGCGCGCGCATTCGATCAGCCATGATGATATCTTTCTCGCAGGAGCAAGGTGAG	4620
Qy	4621	ATGACAGAGATCTCTGCCCGGCACTTCGCCCAATAGCACAGTCCCTTCCCGCTTCAG	4680
Db	4621	ATGACAGAGATCTCTGCCCGGCACTTCGCCCAATAGCACAGTCCCTTCCCGCTTCAG	4680
Qy	4681	TGACAAACGTCGAGCACAGTTCGCAAGGAAGACCCCGCTCGTGGCCAGCCACGATAGCGCG	4740
Db	4681	TGACAAACGTCGAGCACAGTTCGCAAGGAAGACCCCGCTCGTGGCCAGCCACGATAGCGCG	4740
Qy	4741	CTGCCTCGTCTCGATGTTTCATTTCAGGGCACCGGACAGGTCCGTCTTGACAAAAGAACCG	4800
Db	4741	CTGCCTCGTCTCGATGTTTCATTTCAGGGCACCGGACAGGTCCGTCTTGACAAAAGAACCG	4800
Qy	4801	GGCGCCCTCGCTGACAGCGGGAACACCGCGGCGATCAGACACCCGATTTCTGTGTG	4860
Db	4801	GGCGCCCTCGCTGACAGCGGGAACACCGCGGCGATCAGACACCCGATTTCTGTGTG	4860
Qy	4861	CCCAGTCATAGCCGAATAGCTCTCCACCCGAAGCGCGGAGAACCTTGGTCGATCCAT	4920
Db	4861	CCCAGTCATAGCCGAATAGCTCTCCACCCGAAGCGCGGAGAACCTTGGTCGATCCAT	4920
Qy	4921	CTTGTTCATCATGCGAAACCATCCTCATCTGTCTCTTTGATCAGATCTTGATCCCTCTG	4980
Db	4921	CTTGTTCATCATGCGAAACCATCCTCATCTGTCTCTTTGATCAGATCTTGATCCCTCTG	4980
Qy	4981	GCCATCAGATCCTTCGCGGCAAGAACCATCCAGTTTACTTTTCAGGGCTTCCCAACT	5040
Db	4981	GCCATCAGATCCTTCGCGGCAAGAACCATCCAGTTTACTTTTCAGGGCTTCCCAACT	5040

Qy	5041	TACAGAGGCGCCCGCCAGCTGGCAATTC	CGGTTGCGCTTGTCTGTCCATATAAACCGCCAGT	5100			
Db	5041	TACAGAGGCGCCCGCCAGCTGGCAATTC	CGGTTGCGCTTGTCTGTCCATATAAACCGCCAGT	5100			
Qy	5101	CTAGCTATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGGCTTCGCTTTT	5160				
Db	5101	CTAGCTATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGGCTTCGCTTTT	5160				
Qy	5161	CCCTTGTCAGATAGCCCACTAGCTAGCTGACATCATCCGGGGTTCAGACCGGTTTCTGCGGAC	5230				
Db	5161	CCCTTGTCAGATAGCCCACTAGCTAGCTGACATCATCCGGGGTTCAGACCGGTTTCTGCGGAC	5230				
Qy	5221	TGGCTTTCTACGTGTTCCGCTTCCCTTTAGCAGCCCTTTCGCGCCCTGAGTGCTTTGCGGACG	5280				
Db	5221	TGGCTTTCTACGTGTTCCGCTTCCCTTTAGCAGCCCTTTCGCGCCCTGAGTGCTTTGCGGACG	5280				
Qy	5281	GTG 5283					
Db	5281	GTG 5283					
RESULT 2							
AAAI5565							
ID	AAAI5565 standard; DNA; 8349 BP.						
XX	AC	AAAI5565;					
DT	28-JUL-2000 (first entry)						
DE	pMON30464 plasmid.						
XX	Transgenic plant; insect resistance; cry2Aa delta-endotoxin; Coleopteran;						
KW	Lepidopteran; Dipteran; plastid transit peptide; zmsu PTP; insecticidal;						
KW	plastid targeting peptide; pMON30464 plasmid; expression vector; ds.						
OS	Synthetic.						
XX	WC0200026371-Al.						
PD	11-MAY-2000.						
XX	04-NOV-1999; 99WO-US26086.						
PF	04-NOV-1998; 98US-0186002.						
PR	(MONS) MONSANTO CO.						
PA	Corbin DR, Romano CP;						
XX	WPI; 2000-376130/32.						
DR	New method of expressing insecticidal proteins in plants transformed						
PT	with a Bacillus thuringiensis delta-endotoxin encoding gene resulting						
PT	in effective control of susceptible target pests						
XX	Example 1; Page 96-99; 104pp; English.						
PS	Bacillus thuringiensis produce delta-endotoxins during sporulation. These						
XX	proteins are toxic to certain insects e.g. Lepidopteran and Coleopteran						
CC	larvae. An insect-resistant transgenic plant has been constructed which						
CC	contains the delta-endotoxin cry2Ab gene. The present sequence would be						
CC	used to transfer delta-endotoxin genes into plant cells and for						
CC	subsequent high expression of the cry2Ab gene. Protection may be attained						
CC	against insects such as Ostrina spp., Diatraea spp., Helicoverpa spp.,						
CC	and Spodoptera spp., in Zea mays; Heliothis virescens, Helicoverpa spp.,						
CC	Pectinophora spp., in Gossypium hirsutum; Anticarsia spp., Pseudoplusia						
CC	sativa, Epinotia spp., in Glycine max; and Scirpophaga incertulas in Oriza						
CC	protein comprising an amino-terminal plastid transit peptide (zmsu PTP)						
CC	covalently linked to the delta-endotoxin. The fusion protein functions to						
CC	localise the delta-endotoxin to a subcellular organelle or compartment.						

SQ Sequence 8349 BP; 1959 A; 2355 C; 2088 G; 1947 T; 0 other;									
Query Match 42.4%; Score 2240; DB 21; Length 8349;									
Best Local Similarity 93.4%; Pred. No. 0;									
Matches 2409; Conservative 0; Mismatches 25; Indels 144; Gaps 1;									
QY	2706	CTAGAGCTAATCATGTGTCATAGCTGTTTCTCTGTGTAATTTGTTATCCGCTCACAAATCC	2765						
DB	5911	CTTGGGGTAAATCATGTGTCATAGCTGTTTCTCTGTGTAATTTGTTATCCGCTCACAAATCC	5970						
QY	2766	ACACAACTACGAGCGGAGCATAAAGTGTAAAGCTCGGGTGCCTAATGATGAGCTA	2825						
DB	5971	ACACAACTACGAGCGGAGCATAAAGTGTAAAGCTCGGGTGCCTAATGATGAGCTA	6030						
QY	2826	ACTCACATTAATTTGCGTTCGCTCACTGCGCGCTTTCCAGTCGGGAAACCTCTCGTGCCA	2885						
DB	6031	ACTCACATTAATTTGCGTTCGCTCACTGCGCGCTTTCCAGTCGGGAAACCTCTCGTGCCA	6090						
QY	2886	GCTGCATTAAATGTAATCGGCCCAACGCGCGGGAGAGCGGTTTGCTATTTGGCGCTCTTC	2945						
DB	6091	GCTGCATTAAATGTAATCGGCCCAACGCGCGGGAGAGCGGTTTGCTATTTGGCGCTCTTC	6150						
QY	2946	CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	3005						
DB	6151	CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	6210						
QY	3006	TCCTCAAGGCGGTAAATCGGTTATCCACAGAAATCAGGGGTAACGCGAGGAAAGCAAT	3065						
DB	6211	TCCTCAAGGCGGTAAATCGGTTATCCACAGAAATCAGGGGTAACGCGAGGAAAGCAAT	6270						
QY	3066	GTGAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAAGAGCGGCTGCTGGCGTTTTT	3125						
DB	6271	GTGAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAAGAGCGGCTGCTGGCGTTTTT	6330						
QY	3126	CCATAGGCTCCGCCCTTCGAGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCG	3185						
DB	6331	CCATAGGCTCCGCCCTTCGAGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCG	6390						
QY	3186	AAACCCGACAGACTAATAAGATACAGAGGTTTCCCGCTGAGAGCTCCCTCGTGGGCTC	3245						
DB	6391	AAACCCGACAGACTAATAAGATACAGAGGTTTCCCGCTGAGAGCTCCCTCGTGGGCTC	6450						
QY	3246	TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAGGCT	3305						
DB	6451	TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAGGCT	6510						
QY	3306	GGCGCTTTCTCATAGCTCAGCTGTGATAGTATCTCAGTTCCGTTGATGCTGCTCCAA	3365						
DB	6511	GGCGCTTTCTCATAGCTCAGCTGTGATAGTATCTCAGTTCCGTTGATGCTGCTCCAA	6570						
QY	3366	GCTGGGCTGTGTCAGCAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGTTAACTA	3425						
DB	6571	GCTGGGCTGTGTCAGCAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGTTAACTA	6630						
QY	3426	TCGCTTTAGTCCAAACCCCGTAAAGACAGCTTATCCGCACTGGCAGCGCACTGGTAA	3485						
DB	6631	TCGCTTTAGTCCAAACCCCGTAAAGACAGCTTATCCGCACTGGCAGCGCACTGGTAA	6690						
QY	3486	CAGGATTAGCAGCGAGGTATAGCGGTGTCTACAGAGTTCTTGAAGTGTGGCTTAA	3545						
DB	6691	CAGGATTAGCAGCGAGGTATAGCGGTGTCTACAGAGTTCTTGAAGTGTGGCTTAA	6750						
QY	3546	CTACGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTT	3605						
DB	6751	CTACGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTT	6810						
QY	3606	CGGAAAAAGGTTGGTAGCTCTTGTATCCGGCAAAACAAACCGCGCTGGTAGCGGTGTT	3665						
DB	6811	CGGAAAAAGGTTGGTAGCTCTTGTATCCGGCAAAACAAACCGCGCTGGTAGCGGTGTT	6870						
QY	3666	TTTTGTTTGAAGCAGCAGATTAACGGCAGAAAAAAGGATCTCAAGAGATCTCTTTGAT	3725						
DB	6871	TTTTGTTTGAAGCAGCAGATTAACGGCAGAAAAAAGGATCTCAAGAGATCTCTTTGAT	6930						
QY	3726	CTTTTCTACGGGCTCAGCGCTCAGTGAACGAAACTCAGTTTAAGGGATTTTGGTCAT	3785						
DB	6931	CTTTTCTACGGGCTCAGCGCTCAGTGAACGAAACTCAGTTTAAGGGATTTTGGTCAT	6990						
QY	3786	GAGATTATCGTCGACAAAGCGGCATCGTGCCTTCCCACTCTCTGCAAGTTCGGGGGCATG	3845						
DB	6991	GAGATTATCAAAAAGGATCTTTCACCTAGATCTTT-----	7024						
QY	3846	GATCGCGGATAGCGCTGCTGTTTCTCGATGCGGACGGAATTGCACTGCGCGGTAGAA	3905						
DB	7025	-----	7024						
QY	3906	CTCCGCGAGGTCTGTCAGCCTCAGCAGCAGCTGAACAACTCGCGAGGGATCGAGCCC	3965						
DB	7026	-----	7026						
QY	3966	GGGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGGGCTC	4025						
DB	7027	GGGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGGGCTC	7086						
QY	4026	CCGAAAAACGATTCGGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAATCTCG	4085						
DB	7087	CCGAAAAACGATTCGGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAATCTCG	7146						
QY	4086	TGATGGCAGGTTGGGCTCGCTTGGTTCGTCATTTTGAACCCAGAGTCCCGCTCAGAAG	4145						
DB	7147	TGATGGCAGGTTGGGCTCGCTTGGTTCGTCATTTTGAACCCAGAGTCCCGCTCAGAAG	7206						
QY	4146	AACCTGCTCAGAAGCGCATAGAAGCGATGCGTTCGGAATCGGAGCGGGATACCGTAA	4205						
DB	7207	AACCTGCTCAGAAGCGCATAGAAGCGATGCGTTCGGAATCGGAGCGGGATACCGTAA	7266						
QY	4206	AGCAGGAGAGGCTCAGCCCATTTCCGCCCAAGCTTTTTCAGCAATATCACCGGTAGCC	4265						
DB	7267	AGCAGGAGAGGCTCAGCCCATTTCCGCCCAAGCTTTTTCAGCAATATCACCGGTAGCC	7326						
QY	4266	AACGCTATGCTTATAGAGCGTCCGCCACACCCAGCGCGCACAGTCGATGAATCCAGAA	4325						
DB	7327	AACGCTATGCTTATAGAGCGTCCGCCACACCCAGCGCGCACAGTCGATGAATCCAGAA	7386						
QY	4326	AAGCGGCAATTTTCCACCATGATATTCGGCAAGCAGGCATCGCATGGGTCAAGCAGAGA	4385						
DB	7387	AAGCGGCAATTTTCCACCATGATATTCGGCAAGCAGGCATCGCATGGGTCAAGCAGAGA	7446						
QY	4386	TCCTCGCGCTGGGCATCGCGCTTTCAGCCTGGCGAACAGTTTCGGCTGGCGGAGCCCC	4445						
DB	7447	TCCTCGCGCTGGGCATCGCGCTTTCAGCCTGGCGAACAGTTTCGGCTGGCGGAGCCCC	7506						
QY	4446	TGATGCTCTTTCGTCAGATCATCTGATCGAACAGCCGCTTCCATCCGAGTACGTGCT	4505						
DB	7507	TGATGCTCTTTCGTCAGATCATCTGATCGAACAGCCGCTTCCATCCGAGTACGTGCT	7566						
QY	4506	CGCTCGATGCGATGTTTTCGCTTGGTGTGGAATGGGCGAGTACCGGATCAAGCGTATGC	4565						
DB	7567	CGCTCGATGCGATGTTTTCGCTTGGTGTGGAATGGGCGAGTACCGGATCAAGCGTATGC	7626						
QY	4566	AGCCGCGCATTTGCATCAGCCATGATGATCTTCTCGCAGAGCAAGTCAGATGAC	4625						
DB	7627	AGCCGCGCATTTGCATCAGCCATGATGATCTTCTCGCAGAGCAAGTCAGATGAC	7686						
QY	4626	AGGAGATCTTCCCGCGCACTTCGCCAATAGCAGCCAGTCCCTTCCGCTTCAGTGACA	4685						
DB	7687	AGGAGATCTTCCCGCGCACTTCGCCAATAGCAGCCAGTCCCTTCCGCTTCAGTGACA	7746						
QY	4686	ACGTCGAGCAGCTGCGCAGGAAACCGCGCTGTCGCGCAGCCAGTACCGCGGTGCC	4745						
DB	7747	ACGTCGAGCAGCTGCGCAGGAAACCGCGCTGTCGCGCAGCCAGTACCGCGGTGCC	7806						
QY	4746	TCGTCCTTCGAGTTTCAATTCAGGGCACCGGACAGGTTCGCTTTTGACAAAAAGAACCGGGCGC	4805						
DB	7807	TCGTCCTTCGAGTTTCAATTCAGGGCACCGGACAGGTTCGCTTTTGACAAAAAGAACCGGGCGC	7866						

CC the origin of replication and replication stability sequences, and the
 CC promoter and coding sequence for t99 kanamycin/neomycin
 CC phosphotransferase. The plasmid was synthesised from 192 50-mers (see
 CC AAX5201-12) to demonstrate the method of the invention. The
 CC specification describes a method for the synthesis of replication
 CC competent double-stranded polynucleotides. The method comprises
 CC generating a first set of oligonucleotides corresponding to the plus
 CC strand and a second set corresponding to the minus strand and annealing.
 CC The method can be used for preparing polynucleotides encoding sequences
 CC involved in a biochemical pathway. In particular, they can be used to
 CC produce polynucleotides encoding enzymes, e.g. hexokinase, phosphohexose
 CC isomerase, phosphofructokinase-1, aldolase, triose-phosphate isomerase,
 CC glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase,
 CC phosphoglycerate mutase, enolase or pyruvate kinase. They can
 CC also be used for the preparation of viral particles, artificial genomes
 CC and artificial genetic systems.

XX
 SQ Sequence 4800 BP; 1345 A; 1032 C; 1163 G; 1260 T; 0 other;

Query March 33.9%; Score 1793.4; DB 20; Length 4800;
 Best Local Similarity 91.8%; Pred. No. 4.9e-269;
 Matches 1980; Conservative 0; Mismatches 1; Indels 176; Gaps 1;

QY 3127 CATAGCTCCGCCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGCGA 3186
 DB 4800 CATAGCTCCGCCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGCGA 4741
 QY 3187 AACCCGACAGGACTTAAGATACACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCT 3246
 DB 4740 AACCCGACAGGACTTAAGATACACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCT 4681
 QY 3247 CTGTGTTCCGACCTCGCGCTTACCAGGATACCTGTCGCGCTTTCCTTCGGAAGCGTG 3306
 DB 4680 CTGTGTTCCGACCTCGCGCTTACCAGGATACCTGTCGCGCTTTCCTTCGGAAGCGTG 4621
 QY 3307 GCGCTTTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCAAG 3366
 DB 4620 GCGCTTTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCAAG 4561
 QY 3367 CTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACGCTGCGCTTATCCGGTAACTAT 3426
 DB 4560 CTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACGCTGCGCTTATCCGGTAACTAT 4501
 QY 3427 GCTCTTGTAGTCAACCCGCTAGACACACTTTCGCACTGCGACGAGCCACTGGTAC 3486
 DB 4500 GCTCTTGTAGTCAACCCGCTAGACACACTTTCGCACTGCGACGAGCCACTGGTAC 4441
 QY 3487 AGGATTAGCAGCGAGGATGTAGCGGTCTACAGAGTTCTTGAAGTGTGCGCTAAC 3546
 DB 4440 AGGATTAGCAGCGAGGATGTAGCGGTCTACAGAGTTCTTGAAGTGTGCGCTAAC 4381
 QY 3547 TAGCGCTACACTAGAGAACGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTC 3606
 DB 4380 TAGCGCTACACTAGAGAACGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTC 4321
 QY 3607 GGAAGAAGTGTGTAGTCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGTGTTT 3666
 DB 4320 GGAAGAAGTGTGTAGTCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGTGTTT 4261
 QY 3667 TTGTGTTGACGACGACGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGATC 3726
 DB 4260 TTGTGTTGACGACGACGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGATC 4201
 QY 3727 TTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGGTCATG 3786
 DB 4200 TTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGGTCAT - 4142
 QY 3787 AGATTATCTCGAACAAACGCGCCATCGTGCTTCCCACTCTCTGCTGCGGCGCATGG 3846
 DB 4141 ----- 4142
 QY 3847 ATCGCGGATAGCCGCTGTGTTTCTTGATGCGACGAGTTTGCACTGCCGGTAGAAC 3906

DB 4141 ----- 4142
 QY 3907 TCCGCGAGTGTGTCCAGCCTCAGGCAGCAGCTGAACCAACTCGCGAGGGATCGAGCCG 3966
 DB 4141 -----GCCG 4137
 QY 3967 GGTGTGGCGAGNACTCCAGCATGAGATCCCGCGCTGGAGATCATCCAGCGCGCTCC 4026
 DB 4136 GGTGTGGCGAGNACTCCAGCATGAGATCCCGCGCTGGAGATCATCCAGCGCGCTCC 4077
 QY 4027 CGAAAAACGATTCGGAAGCCCAACCTTTATAGAAGCGGGGTGAATCGAAATCTGTT 4086
 DB 4076 CGAAAAACGATTCGGAAGCCCAACCTTTATAGAAGCGGGGTGAATCGAAATCTGTT 4017
 QY 4087 GATGGCAGGTTGGCGCTGCTTGGTCTGCTGATTTTCGAACCCCGAGAGTCCCGCTCAGAAGA 4146
 DB 4016 GATGGCAGGTTGGCGCTGCTTGGTCTGCTGATTTTCGAACCCCGAGAGTCCCGCTCAGAAGA 3957
 QY 4147 ACTGCTCAAGAAGCGCATAGAAGCGATGCGTTCGGAATCGGAGCGCGGATACCGTAAA 4206
 DB 3956 ACTGCTCAAGAAGCGCATAGAAGCGATGCGTTCGGAATCGGAGCGCGGATACCGTAAA 3897
 QY 4207 GCACGAGGAAGCGGTGAGCCCATTTCCCGCCCAAGCTCTTCAGCAATATACCGGTAGCCA 4266
 DB 3896 GCACGAGGAAGCGGTGAGCCCATTTCCCGCCCAAGCTCTTCAGCAATATACCGGTAGCCA 3837
 QY 4267 ACGTATGCTCTGATAGCGTTCGCGCACACCCAGCCGCGCACAGTCTGATGAATTCAGNAA 4326
 DB 3836 ACGTATGCTCTGATAGCGTTCGCGCACACCCAGCCGCGCACAGTCTGATGAATTCAGNAA 3777
 QY 4327 AGGGGCCATTTTCCACCATGATTTTCGGCAAGCAGGATTCGCCATGGGTACACGAGAT 4386
 DB 3776 AGGGGCCATTTTCCACCATGATTTTCGGCAAGCAGGATTCGCCATGGGTACACGAGAT 3717
 QY 4387 CCTCGCGTGGGATCGCGCTTGGAGCTTGGCGAAGCAGTTCGCTGCGCGGAGCCCT 4446
 DB 3716 CCTCGCGTGGGATCGCGCTTGGAGCTTGGCGAAGCAGTTCGCTGCGCGGAGCCCT 3657
 QY 4447 GATGCTTTCGTCAGATCATCTGATCGAAGACCGGCTTCCATCCGAGTACGTGTC 4506
 DB 3656 GATGCTTTCGTCAGATCATCTGATCGAAGACCGGCTTCCATCCGAGTACGTGTC 3597
 QY 4507 GCTCGATGCGATGTTTCGGTGTGATGCGAGTGGCGAGTACGCGATCAAGCGTATGCA 4566
 DB 3596 GCTCGATGCGATGTTTCGGTGTGATGCGAGTGGCGAGTACGCGATCAAGCGTATGCA 3537
 QY 4567 GCGCGCGATGTCATCAGCCATGATGATGATCTTCTCGGAGGAGCAAGGTGAGATGACA 4626
 DB 3536 GCGCGCGATGTCATCAGCCATGATGATGATCTTCTCGGAGGAGCAAGGTGAGATGACA 3477
 QY 4627 GGAGATCTGCGCGGCACTTCGCCCAATAGCAGCAGTTCCTTCCGCTTTCAGTGACAA 4686
 DB 3476 GGAGATCTGCGCGGCACTTCGCCCAATAGCAGCAGTTCCTTCCGCTTTCAGTGACAA 3417
 QY 4687 GTCGAGCAGCTGGCGAAGAAAGCGCGCTGCGCGAGCAGCAGATAGCGCGCTGCT 4746
 DB 3416 GTCGAGCAGCTGGCGAAGAAAGCGCGCTGCGCGAGCAGCAGATAGCGCGCTGCT 3357
 QY 4747 GTCCTGCGAGTTCATTCAGGGCAACCGGACAGTTCGCTTTCGACAAAAGAAACCGGGGCGCC 4806
 DB 3356 GTCCTGCGAGTTCATTCAGGGCAACCGGACAGTTCGCTTTCGACAAAAGAAACCGGGGCGCC 3297
 QY 4807 CCTGCGCTGACCGCGAACAAGCGCGGATCAGACAGCGGATGCTGTTGCTGCCAGT 4866
 DB 3296 CCTGCGCTGACCGCGAACAAGCGCGGATCAGACAGCGGATGCTGTTGCTGCCAGT 3237
 QY 4867 CATAGCGAATAGCTCTCCACCAAGCGCGCGGAGAACTCGCTGCGTGAATCAATCTTGT 4926
 DB 3236 CATAGCGAATAGCTCTCCACCAAGCGCGCGGAGAACTCGCTGCGTGAATCAATCTTGT 3177
 QY 4927 CATATGCGAAGACGATCTCATCTGCTCTTGTATCAGATCTTGTATCCCTCGCGCATC 4986
 DB 3176 CATATGCGAAGACGATCTCATCTGCTCTTGTATCAGATCTTGTATCCCTCGCGCATC 3117

QY 4987 AGATCTTGGCGGCAAGAAAGCCATCCAGTTACTTTGAGGGCTTCCCAACTTACCAG 5046
 Db 3116 AGATCTTGGCGGCAAGAAAGCCATCCAGTTACTTTGAGGGCTTCCCAACTTACCAG 3057
 QY 5047 AGGCGGCCAGCTGGCAATTCGGGTTGCTGCTGCTCCATAAACCCGCGCTAGCT 5106
 Db 3056 AGGCGGCCAGCTGGCAATTCGGGTTGCTGCTGCTCCATAAACCCGCGCTAGCT 2997
 QY 5107 ATCCGCATGTAAGCCCACTGCAAGCTACCTGCTTCTTTGCGTTCGCTTCCCTTG 5166
 Db 2996 ATCCGCATGTAAGCCCACTGCAAGCTACCTGCTTCTTTGCGTTCGCTTCCCTTG 2937
 QY 5167 TCCAGATAGCCAGTACCTGACATTCATCCGGGCTCAGCACCGTTTCTGCGACTGGCTT 5226
 Db 2936 TCCAGATAGCCAGTACCTGACATTCATCCGGGCTCAGCACCGTTTCTGCGACTGGCTT 2877
 QY 5227 TCTAGCTGTTCCGCTTCTTTAGAGCCCTTGGCGCCCTGAGTCTTGGCGAGCGTG 5283
 Db 2876 TCTAGCTGTTCCGCTTCTTTAGAGCCCTTGGCGCCCTGAGTCTTGGCGAGCGTG 2820

RESULT 5

AAV33629/c

ID AAV33629 standard; DNA; 5594 BP.

XX AAV33629;

DT 29-DEC-1998 (first entry)

DE GENSA 981, a monomeric DNA sequence produced by the invention.

KW Class IIS restriction endonuclease recognition site;

KW endogenous mouse promoter element; tissue-specific gene expression;

KW hormone-specific gene expression; as;

KW developmental-specific gene expression.

XX Synthetic.

XX WO9838326-A1.

XX 03-SEP-1998.

XX 28-FEB-1998; 38WO-US03918.

XX 28-FEB-1997; 97US-0070910.

XX (NATU-) NATURE TECHNOLOGY CORP.

XX Hodgson CP, Xu G, Zink MA;

XX WPI; 1998-495399/42.

XX Method for assembling gene or gene vector - comprises use of primers containing

XX Example 2; Pages 107-110; 141pp; English.

XX The invention provides a novel method for directing self-assembly of a gene having three or more fragments in a directionally and spatially ordered fashion to produce a gene or a gene vector. The method involves usage of primers, containing class IIS restriction endonuclease recognition sites, for isolation of these fragments. As described in the disclosure, the method may also use a vector for the incorporation and screening of endogenous mouse promoter elements for the identification of cell specific promoters. In the example given, plasmids pBK-CMV (AAV33626), pVLMB (AAV33623) and pVL0VhGH-900 (AAV33621) were used as templates from which six fragments were amplified. Each of the fragments contained different regulatory sequences. The six PCR fragments were designed to self-assemble into a retro-vector using the method of the invention. The present sequence, designated as GENSA 981, represents the monomeric DNA sequence of the six ligated fragments. In general, the method is claimed to be useful for

CC isolating and identifying regulatory sequences from a cell, including CC those for enhanced biological activity, or tissue-specific, CC hormone-specific or developmental-specific gene expression.

XX Sequence 5594 BP; 1232 A; 1519 C; 1518 G; 1325 T; 0 other;

QY Query Match 31.3%; Score 1652.4; DB 19; Length 5594;

Db Best Local Similarity 85.4%; Pred. No. 3.2e-247;

QY Matches 1916; Conservative 0; Mismatches 261; Indels 67; Gaps 4;

QY 3107 CGCGCTTCTGCGCTTTTCCATAGGCTCCGCCCTTCCAGAGCATCAAAAATCGAC 3166

Db 5591 CGCGCTTCTGCGCTTTTCCATAGGCTCCGCCCTTCCAGAGCATCAAAAATCGAC 5532

QY 3167 GCTCAAGTCAGAGTGGCGAACCAGCAGAGCTATAAGATACCAAGGCTTTCCCTCG 3226

Db 5531 GCTCAAGTCAGAGTGGCGAACCAGCAGAGCTATAAGATACCAAGGCTTTCCCTCG 5472

QY 3227 GAAGCTCCCTCGTGGCTTCTCTGTTCCGACCCCTTACCGGATACCTGTCGCT 3286

Db 5471 GAAGCTCCCTCGTGGCTTCTCTGTTCCGACCCCTTACCGGATACCTGTCGCT 5412

QY 3287 TTCTCCCTTCGCGAAGCGTGGCGCTTTCTCATAGCTCACGCTAGTATCTCAGTTCG 3346

Db 5411 TTCTCCCTTCGCGAAGCGTGGCGCTTTCTCATAGCTCACGCTAGTATCTCAGTTCG 5352

QY 3347 TGTAGTCTGCTTCCAGCTGGGCTGTGTGACAAACCCCGCTTACGCCGAGCGCT 3406

Db 5351 TGTAGTCTGCTTCCAGCTGGGCTGTGTGACAAACCCCGCTTACGCCGAGCGCT 5292

QY 3407 GCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCGCTAAGACAGACTTATCCGCAC 3466

Db 5291 GCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCGCTAAGACAGACTTATCCGCAC 5232

QY 3467 TGGCAGCAGCCACTGCTTAACAGATTAGCAGAGGATGTAGCGGCTGTACAGAT 3526

Db 5231 TGGCAGCAGCCACTGCTTAACAGATTAGCAGAGGATGTAGCGGCTGTACAGAT 5172

QY 3527 TCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGGGCTC 3586

Db 5171 TCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGGGCTC 5112

QY 3587 TGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGCAACCAACCA 3646

Db 5111 TGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGCAACCAACCA 5052

QY 3647 CCGCTGTAGCGGTGTTTCTTTTCTTTCGAGCAGCAGATTACCGCAGAAAAAGGAT 3706

Db 5051 CCGCTGTAGCGGTGTTTCTTTTCTTTCGAGCAGCAGATTACCGCAGAAAAAGGAT 4992

QY 3707 CTCAGAAGATCCCTTTGATCTTTTCTACGGGCTGTGACGCTC-----AGTGGACG 3757

Db 4991 GGTGTGGGCTCTTTTATTGAGCTCGGGAGCAGACGCGCGAAGCAGAGAGCG 4932

QY 3758 AAAACTCAGTTAAGGGATTTTGTCTATGAGATTATCGTCGACCAAGCGGCGCATCGTGC 3817

Db 4931 AACTGATTGGTTAGTTCAATAAGGCACAGGTCATTTACAGTCTCTTGGGCGACCTTGA 4872

QY 3818 CTC-----CCCACTCTCTGAGTTCCGGGCGCATGGATCGCGGATAGCGGCTG 3864

Db 4871 AACATCTGATGGTCTCTTAGAAAATGCTGAGGCTGGACCGCATCTGGGAGCCATCTGTT 4812

QY 3865 CTGGTTTCTGATGCCGACCGGATTTG-----ATGCGCGTAGAACCTCCGCGAG 3914

Db 4811 CTGGCGCTTGAAGCGCGGCGCAGGAACTGCTTACCACAGATATCTGTTTGGCCCATATTC 4752

QY 3915 GTCTGTCAGCTCAGGCGCAGAGCTGAACCAACTCGCAGGGGATCGAGCCCGGGGTGGG 3974

Db 4751 GTCTGTCAGCTCAGGCGCAGAGCTGAACCAACTCGCAGGGGATCGAGCCCGGGGTGGG 4692

QY 3975 GAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCCGCGCTCCCGGAAAC 4034

Db 4691 TTGGCCCATATTCAGCTGTTCATCTCTGTTCTGACCTTGTGATCTGAACTTCTTATCTC 4632

QY 4035 GATTCGAGCCCAACCTTTTCATAGAGCGCGGTGGAATCGAAATCTCG----- 4085
DB 4631 AGTTATGATATTTTCCATGCTTGCATAAATGGGTTACTTAAGCTAGCTTGCCAAACCTA 4572
QY 4086 -----TGATGGCAGGTTGGCGCTCGCTTGGTGGCTGCTCAT 4119
DB 4571 CAGTGGGGTCTTTCAATCCGCCCTTTTCTGGAGGTTGGCGTCTGCTTGGTGGCTCAT 4512
QY 4120 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGCATAGAGGCGATCGCT 4179
DB 4511 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGCATAGAGGCGATCGCT 4452
QY 4180 GCGATCCGGAGCGCGATACCGTAAAGCAGGAGCGGTGAGCCCATTCGCGCCAA 4239
DB 4451 GCGATCCGGAGCGCGATACCGTAAAGCAGGAGCGGTGAGCCCATTCGCGCCAA 4392
QY 4240 GCTCTTCAGCAATATCAGCGTAGCCAAAGCTATGCTCTGATAGCGGTGCGCCACACCA 4299
DB 4391 GCTCTTCAGCAATATCAGCGTAGCCAAAGCTATGCTCTGATAGCGGTGCGCCACACCA 4332
QY 4300 GCGGCCACAGTCGATGATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAGC 4359
DB 4331 GCGGCCACAGTCGATGATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAGC 4272
QY 4360 AGGCATCCCATGGTTCAGCAGAGATCTCGCGTCGGGATCGCGCTTGAGCTGG 4419
DB 4271 AGGCATCCCATGGTTCAGCAGAGATCTCGCGTCGGGATCGCGCTTGAGCTGG 4212
QY 4420 CGAACAGTTGCGTGGCGGAGCCCTGATGCTCTGTCAGATCATCTGATGACAA 4479
DB 4211 CGAACAGTTGCGTGGCGGAGCCCTGATGCTCTGTCAGATCATCTGATGACAA 4152
QY 4480 GACCGCTTCATCCGATAGCTGCTGCTCGATGCGATGTTGCTTGGTGGTGAATG 4539
DB 4151 GACCGCTTCATCCGATAGCTGCTGCTCGATGCGATGTTGCTTGGTGGTGAATG 4092
QY 4540 GCGAGTAGCGGATCAAGCGGTATGACGCCCGCATTTGATCAGCCATGATGATCTT 4599
DB 4091 GCGAGTAGCGGATCAAGCGGTATGACGCCCGCATTTGATCAGCCATGATGATCTT 4032
QY 4600 TCTGCGAGGAGCAAGGTGAGATGACAGAGATCTCGCGCGGACCTTCGCCCAATAGCA 4659
DB 4031 TCTGCGAGGAGCAAGGTGAGATGACAGAGATCTCGCGCGGACCTTCGCCCAATAGCA 3972
QY 4660 GCGAGTCCCTTCCGCTTCAGTGACAGCTGACAGAGCTGCGCAAGGAAACCGCGTGC 4719
DB 3971 GCGAGTCCCTTCCGCTTCAGTGACAGCTGACAGAGCTGCGCAAGGAAACCGCGTGC 3912
QY 4720 TGGCCAGCCAGATAGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4779
DB 3911 TGGCCAGCCAGATAGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3952
QY 4780 CGGTCTTGACAAAAGAACCGCGCGCTTCCGCTGACAGCGGAAACCGCGCGCATCAG 4839
DB 3851 CGGTCTTGACAAAAGAACCGCGCGCTTCCGCTGACAGCGGAAACCGCGCGCATCAG 3792
QY 4840 AGCAGCGGATGCTGTTGTCGCGCTGATAGCGGATAGCTTCCACCGAGCGCGCG 4899
DB 3791 AGCAGCGGATGCTGTTGTCGCGCTGATAGCGGATAGCTTCCACCGAGCGCGCG 3732
QY 4900 GAGAACCTCGTGCATCATCTTGTTCATATCATGCGAAACGATCCTCATCTGCTCTT 4959
DB 3731 GAGAACCTCGTGCATCATCTTGTTCATATCATGCGAAACGATCCTCATCTGCTCTT 3672
QY 4960 GATCAGATCTGATCCCTGCGCATCATCTTGGCGGCAAGAACCATCAGTTTA 5019
DB 3671 GATCAGATCTGATCCCTGCGCATCATCTTGGCGGCAAGAACCATCAGTTTA 3612
QY 5020 CTTTGCAGGCTTCCCAACTTACAGAGGCGCGCCAGCTGCGCAATTCGGTTCGCTTG 5079
DB 3611 CTTTGCAGGCTTCCCAACTTACAGAGGCGCGCCAGCTGCGCAATTCGGTTCGCTTG 3552

QY 5080 CTCTCCATAAAGCCCGCCAGTCTAGCTATCGCATGTAAAGCCACTGCAAGCTACCTCT 5139
DB 3551 CTCTCCATAAAGCCCGCCAGTCTAGCTATCGCATGTAAAGCCACTGCAAGCTACCTCT 3492
QY 5140 TTCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 5199
DB 3491 TTCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3432
QY 5200 GTGAGCAGCGTTTCTCGGACTGGCTTCTAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 5259
DB 3431 GTGAGCAGCGTTTCTCGGACTGGCTTCTAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3372
QY 5260 GCCTGAGTCTTGGCGGAGCGGTG 5283
DB 3371 GCCTGAGTCTTGGCGGAGCGGTG 3348

RESULT 6
AAH74865/c
ID AAH74865 standard; DNA; 5285 BP.
XX AAH74865;
XX
XX 29-OCT-2001 (first entry)
XX Nucleotide sequence of a construct comprising the PCV Rep gene.
XX Rolling circle replication; RCR; Rep gene; gene function; Gemminivirus;
XX Circovirus; Nanovirus; gene therapy; PCV; ss.
XX Synthetic.
XX Porcine circovirus.
XX WO200161024-A2.
XX 23-AUG-2001.
XX 15-FEB-2001; 2001WO-US05394.
XX 16-FEB-2000; 2000US-0505477.
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX Palmer KE, Pogue GP;
XX WPI; 2001-522601/57.
XX New polynucleotide capable of rolling circle replication in host.
XX useful for discovery of gene function, comprises Rep gene, sequences
XX that are cis on the polynucleotide and open reading frame encoding
XX protein of interest
XX Example 1; Fig 5; 61pp; English.
XX The specification describes a polynucleotide that is capable of rolling
XX circle replication (RCR) in an eukaryotic host. The polynucleotide
XX comprises a Rep gene encoding from a virus, sequences that are cis on
XX the polynucleotide such that the Rep protein can bring about RCR of the
XX polynucleotide, an open reading frame encoding a protein of interest
XX capable of being expressed in the host, and optionally a multiple
XX cloning site, where the polynucleotide lacks genes of the virus. The
XX virus is selected from the group of genera of family geminiviridae,
XX circoviridae and nanovirus. The polynucleotide sequence is useful for
XX the discovery of the function of genes in eukaryotic cells, and for
XX inducing or enhancing a trait in a host eukaryotic cell, for
XX down-regulating a gene in a plant or in mammalian cell and thus altering
XX or even eliminating the function of that gene, as gene sequence delivery
XX tools for mammalian genomic approaches, for gene therapy applications
XX (for delivering therapeutic or complementing gene products to organisms
XX or cells), and in whole animal genomics. The present sequence represents
XX a construct comprising the whole PCV genome. This construct contains the
XX PCV Rep gene under the transcriptional control of its own promoter, and
XX has the putative coat protein inactivated by insertion of the bacterial

Db 1695 ----- 1696
 Qy 3861 GCTGCTGTTTCTTGATGCGGAGGATTTGCACTGCGGTAGAACTCCGGAGGTCGTC 3920
 Db 1695 ----- 1696
 Qy 3921 CAGCCTCAGGACAGCTGAACCACTCCGAGGGGATCGAGCCGGGGTGGGGAAGAA 3980
 Db 1695 ----- 1696
 Qy 3981 CTCAGCATGAGATCCCGCGCTGAGAGATCATCAGCGCGGGTCCCGGAAACGATTC 4040
 Db 1674 CTCAGCATGAGATCCCGCGCTGAGAGATCATCAGCGCGGGTCCCGGAAACGATTC 1615
 Qy 4041 GAAGCCCAACTTTCATAGAGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGG 4100
 Db 1614 GAAGCCCAACTTTCATAGAGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGG 1555
 Qy 4101 CGTCGCTTGCTGCTGATTTTCGAAACCCAGAGTCCGCTCAGAGAACTCGTCAAGAGG 4160
 Db 1554 CGTCGCTTGCTGCTGATTTTCGAAACCCAGAGTCCGCTCAGAGAACTCGTCAAGAGG 1495
 Qy 4161 CGATAGAGGCGATGCGCTGCGAATCGGAGCGGGATACCGTAAAGCAGGAGAGCGG 4220
 Db 1494 CGATAGAGGCGATGCGCTGCGAATCGGAGCGGGATACCGTAAAGCAGGAGAGCGG 1435
 Qy 4221 TCAGCCCATTTGCGCGCGAGCTCTTCAGCAATATCAGGGTAGCCAGCTATGCTCGA 4280
 Db 1434 TCAGCCCATTTGCGCGCGAGCTCTTCAGCAATATCAGGGTAGCCAGCTATGCTCGA 1375
 Qy 4281 TAGCGGTCCGCCACACCCAGCGGGCCACAGTCCGATGAATCCAGAAAGCGGCCATTTTC 4340
 Db 1374 TAGCGGTCCGCCACACCCAGCGGGCCACAGTCCGATGAATCCAGAAAGCGGCCATTTTC 1315
 Qy 4341 ACCATGATATTCGCGAGGAGGATCGCATCGGTGATCAGAGAGATCTCGCGCTCGGC 4400
 Db 1314 ACCATGATATTCGCGAGGAGGATCGCATCGGTGATCAGAGAGATCTCGCGCTCGGC 1255
 Qy 4401 ATGCGCGCTTGAGCTCGGAGAGTTCGCTGCGGCGAGCCCTGATGCTCTTCGTC 4460
 Db 1254 ATGCGCGCTTGAGCTCGGAGAGTTCGCTGCGGCGAGCCCTGATGCTCTTCGTC 1195
 Qy 4461 AGATCATCTGATCGACAGACCGGCTTCATCCGAGTACGTCGCTCCGATCGATGT 4520
 Db 1194 AGATCATCTGATCGACAGACCGGCTTCATCCGAGTACGTCGCTCCGATCGATGT 1135
 Qy 4521 TTGCTTTGGTGGATGCGATGCGGAGGATGCGGATCAAGCGTATGAGCGCGCGATTGCA 4580
 Db 1134 TTGCTTTGGTGGATGCGATGCGGAGGATGCGGATCAAGCGTATGAGCGCGCGATTGCA 1075
 Qy 4581 TCAGCCATGATGATATCTTTCTCGGAGGAGCAAGTGAAGTACAGAGATCTTCGCC 4640
 Db 1074 TCAGCCATGATGATATCTTTCTCGGAGGAGCAAGTGAAGTACAGAGATCTTCGCC 1015
 Qy 4641 GGCACTTCGCCAATAGACGCAAGTCCCTTCGCGCTTCAGTGACAACTGCGAGCAGCT 4700
 Db 1014 GGCACTTCGCCAATAGACGCAAGTCCCTTCGCGCTTCAGTGACAACTGCGAGCAGCT 955
 Qy 4701 GCGCAAGGAACCGCGCTGCGGAGCAGCAGATAGCGCGGCTCGCTCTTCGAGTTCA 4760
 Db 954 GCGCAAGGAACCGCGCTGCGGAGCAGCAGATAGCGCGGCTCGCTCTTCGAGTTCA 895
 Qy 4761 TTCAGGCGACGCGAGTCTGCTTTCGAAAGAACCGGGCGGCTTCGCTGACAGC 4820
 Db 894 TTCAGGCGACGCGAGTCTGCTTTCGAAAGAACCGGGCGGCTTCGCTGACAGC 835
 Qy 4821 CCGAACCGCGGATCAGAGCAGCGGATTTGCTGTTGCGGCTCATAGCCGATAGC 4880
 Db 834 CCGAACCGCGGATCAGAGCAGCGGATTTGCTGTTGCGGCTCATAGCCGATAGC 775
 Qy 4881 CTCCTCCACCGAGCGCGGAGACTGCTGCGAATCTCTGTCATCATCGGAAC 4940
 Db 774 CTCCTCCACCGAGCGCGGAGACTGCTGCGAATCTCTGTCATCATCGGAAC 715

RESULT 8

ABQ78294

ID ABQ78294 standard; DNA; 2192 BP.

XX AC ABQ78294;

XX AC ABQ78294;

XX 05-NOV-2002 (first entry)

XX Nucleotide sequence of plasmid pVC0289.

XX Pig; growth hormone releasing hormone; GHRH; growth rate;

XX lean body mass; insulin-like Growth Factor-I; IGF-I; milk production;

XX feed efficiency; somatotroph; growth-related disorder;

XX hypopituitary dwarfism; meat production; egg production; ss.

XX Synthetic.

XX WO200261037-A2.

XX 08-AUG-2002.

XX 12-DEC-2001; 2001WO-US48726.

XX 12-DEC-2000; 2000US-255021P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX (ADVI-) ADVISYS.

XX Schwartz RJ, Carpenter RH, Draghia-Akli R, Kern DR, Smith RG;

XX WPI; 2002-619237/66.

XX Improving or enhancing growth, lean body mass, milk production, feed

XX efficiency or insulin-like Growth Factor-I levels, comprises

XX introducing a vector encoding a growth hormone releasing hormone into

XX an animal before or during gestation -

XX Example 2; Page 111-112; 113pp; English.

XX The present sequence represents plasmid pVC0289, which was used to

XX create a vector carrying cDNA encoding pig growth hormone releasing

XX hormone (GHRH) cDNA. Nucleic acids encoding GHRH are used in the

XX method of the invention. The specification describes a method for

XX improving or enhancing characteristics e.g. growth, lean body mass,

XX insulin-like Growth Factor (IGF)-I levels, growth rate and milk

XX production in an offspring, and for delaying birth of an offspring.

XX The method comprises introducing a vector, encoding GHRH, into cells

CC of the female animal prior to or during gestation of the offspring
CC under conditions where the nucleotide sequence is expressed. The
CC method is useful of improving or enhancing animal growth, for
CC increasing growth hormone, lean body mass, IGF-I levels, feed
CC efficiency, growth rate, ratio of somatostrophs to other
CC hormone-producing cells in a pituitary gland, and milk production in
CC an offspring, and for delaying birth of an offspring. GHRH nucleic
CC acids and vectors are used for diagnostic purposes in clinical
CC medicine, both human and veterinary, e.g. in treating growth-related
CC disorders such as hypopituitary dwarfism resulting from abnormalities
CC in growth hormone production, and in stimulating the growth and
CC enhancing feed conversion efficiency of animals raised for meat, milk
CC and egg production.
XX
SQ Sequence 2192 BP; 486 A; 634 C; 601 G; 471 T; 0 other;

Query Match 30.4%; Score 1603.8; DB 24; Length 2192;
Best Local Similarity 82.8%; Pred. No. 1.2e-239;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;

Qy 2706 CTAGAGGTATATCATGCTATAGCTGTTCTCTGTGTGAATTTATCCGCTCAATTC 2765
Db 70 CTGGGGTAATCATGCTATAGCTGTTCTCTGTGTGAATTTATCCGCTCAATTC 129
Qy 2766 ACAACAATACGAGCGGAGCATAAAGTAAAGCTGGGGTCCCTAATGAGTGAGCTA 2825
Db 130 ACACACATACGAGCGGAGCATAAAGTAAAGCTGGGGTCCCTAATGAGTGAGCTA 189
Qy 2826 ACTCACATTAATGCGTTGCGCTACTGCGCGCTTTCCAGTCGGGAAACCTGCTGGCCA 2885
Db 190 ACTCACATTAATGCGTTGCGCTACTGCGCGCTTTCCAGTCGGGAAACCTGCTGGCCA 249
Qy 2886 GCTGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATTCGGGCGCTCTC 2945
Db 250 GCTGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATTCGGGCGCTCTC 309
Qy 2946 CGTTCTCTGCTCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3005
Db 310 CGTTCTCTGCTCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
Qy 3006 TCATCAAGGCGGTAAATAGGTTATCCACAGAAATCAGGGGATAACCGAGGAAGAATAT 3065
Db 370 TCATCAAGGCGGTAAATAGGTTATCCACAGAAATCAGGGGATAACCGAGGAAGAATAT 429
Qy 3066 GTGAGCAAAAGGCGAGCAAAAGGCGAGGAAACCGTAAAGGCGCGTTGCTGGCGTTTTT 3125
Db 430 GTGAGCAAAAGGCGAGCAAAAGGCGAGGAAACCGTAAAGGCGCGTTGCTGGCGTTTTT 489
Qy 3126 CCAATAGCTCCGCCCCCTGAGAGGATCAAAAATCAGCTCAAGTCAAGAGTGGCG 3185
Db 490 CCAATAGCTCCGCCCCCTGAGAGGATCAAAAATCAGCTCAAGTCAAGAGTGGCG 549
Qy 3186 AAACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTC 3245
Db 550 AAACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTC 609
Qy 3246 TCCTGTTCGACCTCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAGCGT 3305
Db 610 TCCTGTTCGACCTCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAGCGT 669
Qy 3306 GGGCGTTTCTATAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3365
Db 670 GGGCGTTTCTATAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 729
Qy 3366 GCTGGGCTGTGTGTCAGCAACCCCGTTTCCAGCGCGCTGCGCTTATCCGTTAACTA 3425
Db 730 GCTGGGCTGTGTGTCAGCAACCCCGTTTCCAGCGCGCTGCGCTTATCCGTTAACTA 789
Qy 3426 TCGTCTTGAAGTCAACCCCGTTAAGACAGCACTTATCGCACTGACAGACCACTGGTAA 3485
Db 790 TCGTCTTGAAGTCAACCCCGTTAAGACAGCACTTATCGCACTGACAGACCACTGGTAA 849
Qy 3486 CAGGATTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA 3545

Db 850 CAGGATTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA 909
Qy 3546 CTACGGCTACATAGAGCAAGATATTTGGTATCTGGGCTCTGCTGAGCAGCTTACCTT 3605
Db 910 CTACGGCTACATAGAGCAAGATATTTGGTATCTGGGCTCTGCTGAGCAGCTTACCTT 969
Qy 3606 CGAAAAAGAGTTGGTGTAGCTTTGTATCGGCAAAACCAACCGCTGTTAGCGGTGTTT 3665
Db 970 CGAAAAAGAGTTGGTGTAGCTTTGTATCGGCAAAACCAACCGCTGTTAGCGGTGTTT 1029
Qy 3666 TTTTGTGTTCAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAATCTTTGAT 3725
Db 1030 TTTTGTGTTCAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAATCTTTGAT 1089
Qy 3726 CTCTTCTACGGGCTGTCAGCTCAGTGGAAACGAAATCTCAGTTAAGGATTTTGTCTAT 3785
Db 1090 CTCTTCTACGGGCTCTGA----- 1107
Qy 3786 GAGATTATCTCGACCAAGGGCCCATCGTCCCTCCCACTCTCTGCAAGTTTCGGGGGATG 3845
Db 1108 ----- 1107
Qy 3846 GATGCGCGATAGCCGCTGCTGGTTTCTGATGCGGACGATTTGCACTGCCGTAGAA 3905
Db 1108 ----- 1107
Qy 3906 CTCCGCGAGGTCGTCAGGCTCAGCAGCAGCTGAAACAACTCCGAGGGGATCGAGCCC 3965
Db 1108 ----- 1107
Qy 3966 GGGGTGGCGAAGAACTCCAGCATGAGATCCCGCTGGAGGATCATCCAGCGCGCTC 4025
Db 1108 ----- 1107
Qy 4026 CCGAAAAACGATTCGGAAGCCCAACCTTTATAGAGGCGGCGGTGGAATCGAATCTCG 4085
Db 1108 ----- 1107
Qy 4086 TGATGCGAGTTGGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4145
Db 1108 -----CGTTCAGAAG 1117
Qy 4146 AACTCGTCAAGAGCGATAGAGCGGATGCGTTCGGAATCGGGAGCGCGATACCGTAA 4205
Db 1118 AACTCGTCAAGAGCGATAGAGCGGATGCGTTCGGAATCGGGAGCGCGATACCGTAA 1177
Qy 4206 AGCAGGAGAGCGGTTCAGCCCTTCGCGCGCAAGCTTTTCAGCAATATCACGGGTAGCC 4265
Db 1178 AGCAGGAGAGCGGTTCAGCCCTTCGCGCGCAAGCTTTTCAGCAATATCACGGGTAGCC 1237
Qy 4266 AAGCTATCTCTGTATAGCGTCCGCCACACCCAGCGCGCCACAGTCAATTCAGAA 4325
Db 1238 AAGCTATCTCTGTATAGCGTCCGCCACACCCAGCGCGCCACAGTCAATTCAGAA 1297
Qy 4326 AAGCGGCCATTTTCCACCATGATATTCGGAAGCAGACATCGCCATGGTTCAGCAGAGA 4385
Db 1298 AAGCGGCCATTTTCCACCATGATATTCGGAAGCAGACATCGCCATGGTTCAGCAGAGA 1357
Qy 4386 TCTTCGCGCTTCGGGATGCGCGCTTCAGCTGCGCAACAGTTTCGCTGCGCGGAGCCCC 4445
Db 1358 TCTTCGCGCTTCGGGATGCGCGCTTCAGCTGCGCAACAGTTTCGCTGCGCGGAGCCCC 1417
Qy 4446 TGATGCTCTTCGTCAGATCATCTGATCGCAAGACCGGCTTCATCCAGTACGTTGCT 4505
Db 1418 TGATGCTCTTCGTCAGATCATCTGATCGCAAGACCGGCTTCATCCAGTACGTTGCT 1477
Qy 4506 CGCTCATGCGATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4565
Db 1478 CGCTCATGCGATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1537
Qy 4566 AGCGCGCGCATTCGATCGACCATGATGATGATCTTCTCGCGAGGAGCAAGGTGAGTAC 4625

Db 2012 GCGCTTTCTCATAGCTCAGCTGTAGTATCTCAGTTGCGTGTAGTGTGCTGCTCAA 2071
Qy 3366 GCTGGGCTGTGTGACGAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGGTAACTA 3425
Db 2072 GCTGGGCTGTGTGACGAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGGTAACTA 2131
Qy 3426 TCGTCTTGTAGTCCAAACCCGCTAGACAGACTATTCGCCACTTGGCAGCAGCAGTGTAA 3485
Db 2132 TCGTCTTGTAGTCCAAACCCGCTAGACAGACTATTCGCCACTTGGCAGCAGCAGTGTAA 2191
Qy 3486 CAGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTTCAAGTGTGTGCTTAA 3545
Db 2192 CAGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTTCAAGTGTGTGCTTAA 2251
Qy 3546 CTAGCGGTACACTAGAGAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 3605
Db 2252 CTAGCGGTACACTAGAGAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 2311
Qy 3606 CGGAAAGAGTTGCTAGCTTTGATCCGGCAACAAACCAACCGCTGTGTAGCGGTGTTT 3665
Db 2312 CGGAAAGAGTTGCTAGCTTTGATCCGGCAACAAACCAACCGCTGTGTAGCGGTGTTT 2371
Qy 3666 TTTTGTGTTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAAATCTTTGAT 3725
Db 2372 TTTTGTGTTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAAATCTTTGAT 2431
Qy 3726 TTTTGTGTTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAAATCTTTGAT 3785
Db 2432 TTTTGTGTTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAAATCTTTGAT 2449
Qy 3786 GAGATTATCTGTCGACAAAGCGGCAATCGTCCCTCCCTCCTGCTGCGGCGCATG 3845
Db 2450 ----- 2449
Qy 3846 GATCGCGGATAGCGGCTGTGTTTCTGATGCGGACGAGTTTGCATCTGCGGTAGAA 3905
Db 2450 ----- 2449
Qy 3906 CTCCGCGAGGTGCTCCAGCTTCCAGCAGCAGCTGAACCACTCGCGAGGGATCGAGCC 3965
Db 2450 ----- 2449
Qy 3966 GGGGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCCGCGTC 4025
Db 2450 ----- 2449
Qy 4026 CCGGAAACGATTCCGAAGCCCAACCTTTATAGAAAGCGCGGTGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
Qy 4086 TGATGCGAGTTGGCGCTGCTTGTGCTCATTTTCGAACCCCGAGAGTCCCGCTCAGAAG 4145
Db 2450 ----- 2459
Qy 4146 AACTGCTCAAGAGCGATAGAGCGATGCGGTGCGAATCGGAGCGCGATACCGTAA 4205
Db 2460 AACTGCTCAAGAGCGATAGAGCGATGCGGTGCGAATCGGAGCGCGATACCGTAA 2519
Qy 4206 AGCAGAGAGCGGTGAGCGCATTCGCGCCCAAGCTCTTCAGCAATATACCGGTAGCC 4265
Db 2520 AGCAGAGAGCGGTGAGCGCATTCGCGCCCAAGCTCTTCAGCAATATACCGGTAGCC 2579
Qy 4266 AAGCTATGCTCTGATAGGTTGCGCACACCGCGCGCAGCTGATGATCCAGAA 4325
Db 2580 AAGCTATGCTCTGATAGGTTGCGCACACCGCGCGCAGCTGATGATCCAGAA 2639
Qy 4326 AAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTGACGAGAA 4385
Db 2640 AAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTGACGAGAA 2699
Qy 4386 TCTGCGCGTGGGATCGGCTTGAAGCTGCGGACAGTTGCGTGGCGGAGCC 4445

Db 2700 TCCTCGCGCTCGGCGATGCGCGCTTGAGCCTGGCGAACAGTTTCGGCTGGCGCGAGCCCC 2759
Qy 4446 TGATGCTTTGCTCCAGATCATCTGATCGAACAGACCGGCTTCCATCCGAGTACGTGCT 4505
Db 2760 TGATGCTTTGCTCCAGATCATCTGATCGAACAGACCGGCTTCCATCCGAGTACGTGCT 2819
Qy 4506 CGCTCGATGCGATGTTTTCGTTGGTGTGGAATGGCAGGTAGCCGGATCAAGAGCGTATGC 4565
Db 2820 CGCTCGATGCGATGTTTTCGTTGGTGTGGAATGGCAGGTAGCCGGATCAAGAGCGTATGC 2879
Qy 4566 AGCCGCGCATTTGATCATGAGCCATGATGATTAATTTCTCGCAGCAGCAAGGTGATGATAC 4625
Db 2880 AGCCGCGCATTTGATCATGAGCCATGATGATTAATTTCTCGCAGCAGCAAGGTGATGATAC 2939
Qy 4626 AGGAGATCTCTCCCGCGCACTTTCGCCCAATAGCAGCAGTTCCTTCCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTCTCCCGCGCACTTTCGCCCAATAGCAGCAGTTCCTTCCCGCTTCAGTGACA 2999
Qy 4686 AGCTCGACACAGCTGCGCAGGAAACCGCGCTGTCGCGCAGCCACGATAGCCGCTGCC 4745
Db 3000 AGCTCGACACAGCTGCGCAGGAAACCGCGCTGTCGCGCAGCCACGATAGCCGCTGCC 3059
Qy 4746 TCGTCTCTCAGTTCAATTCAGGCGCACCGGACAGCTCGGTCTTGACAAAAAGAACCGGCGC 4805
Db 3060 TCGTCTCTCAGTTCAATTCAGGCGCACCGGACAGCTCGGTCTTGACAAAAAGAACCGGCGC 3119
Qy 4806 CCTCGGCTGACCGCGAACACCGCGGCTCAGAGCAGCCGATTCGTTGTGCGCCAG 4865
Db 3120 CCTCGGCTGACCGCGAACACCGCGGCTCAGAGCAGCCGATTCGTTGTGCGCCAG 3179
Qy 4866 TCATAGCGAATAGCTTCTCCACCAAGCGCGGAGAACCTCGGTGCAATCCATCTTGT 4925
Db 3180 TCATAGCGAATAGCTTCTCCACCAAGCGCGGAGAACCTCGGTGCAATCCATCTTGT 3239
Qy 4926 TCAATCATGCGAAGCAGTCTCATCTGCTCTTGATCAGATCTTGATCCCTTCGCGCAT 4985
Db 3240 TCAATCATGCGAAGCAGTCTCATCTGCTCTTGATCAGATCTTGATCCCTTCGCGCAT 3299
Qy 4986 CAGATCTTTGGCGCAGAAAGCCATTCAGTTTACTTTGAGGCTTCCCACTTACCA 5045
Db 3300 CAGATCTTTGGCGCAGAAAGCCATTCAGTTTACTTTGAGGCTTCCCACTTACCA 3359
Qy 5046 GAGGCGCGCCAGCTGCGAATTCGCGTTGCTGCTGCTCCATAAAAACCGCCAGTGTAGC 5105
Db 3360 GAGGCGCGCCAGCTGCGAATTCGCGTTGCTGCTGCTCCATAAAAACCGCCAGTGTAGC 3419
Qy 5106 TATCGCATGTAAAGCCCACTGCAAGCTACCTGCTTCTTCTT 5146
Db 3420 AACTGTGGGAAGGCGATCGTGGCGGCTCTTCGCTATT 3460

RESULT 10
AAV50428
ID AAV50428 standard; DNA; 3599 bp.
XX AAV50428;
AC AAV50428;
XX 26-OCT-1998 (first entry)
DT 26-OCT-1998 (first entry)
XX
DE Plasmid pIG0552 lower sequence nucleotides 1-3599.
XX Human; IGF-1; insulin-like growth factor 1; urinary incontinence;
KW gene therapy; neurotrophic factor; ss.
XX Synthetic.
OS Homo sapiens.
XX
FN W0983529-A1.
XX
PD 06-AUG-1998.
XX
PF 04-FEB-1998; 98WO-US02051.
XX

PR 04-FEB-1997; 97US-0036862.
 XX (GENE-) GENEMEDICINE INC.
 XX Coleman M;
 XX WPI; 1998-437184/37.
 XX Treatment of urinary incontinence - by delivering nucleic acid
 XX vector for expression of growth factor or neurotrophic factor in
 XX tissue(s)
 XX Disclosure; Page 105-107; 117pp; English.
 XX A method has been developed of treating urinary incontinence (UI) in
 XX mammals. The method comprises delivering a nucleic acid vector for the
 XX expression of a growth factor or neurotrophic factor in a tissue or
 XX tissues. The present sequence represents the actual lower sequence
 XX of plasmid pIG0552 nucleotides 1-3599 from the present invention. Due
 XX to the growth and stimulatory effects of growth factors and neurotrophic
 XX factors, introducing these factors to degenerated muscles in the
 XX urinary system can improve UI by enhancing both their integrity and
 XX neural innervation.
 XX Sequence 3599 BP; 733 A; 1098 C; 1042 G; 726 T; 0 other;
 SQ Query Match 30.4%; Score 1603.8; DB 19; Length 3599;
 Best Local Similarity 82.8%; Pred. No. 1.1e-239;
 Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;

Qy 2706 CTGAGCATATCATGTGCTAGCTGTTCTCTGTGGAATTTGTTATCGCTCACAATTC 2765
 Db |||||
 Qy 56 CTTGGCGTAATCATGTGCTAGCTGTTCTCTGTGGAATTTGTTATCGCTCACAATTC 115
 Db |||||
 Qy 2766 ACACAAATACGAGCGCGGAAGATAAGTAAAGCTTGGGGTGCCTAATGAGTAGCTA 2825
 Db |||||
 Qy 116 ACACAAATACGAGCGCGGAAGATAAGTAAAGCTTGGGGTGCCTAATGAGTAGCTA 175
 Db |||||
 Qy 2826 ACTCAGATTAATTCGTTGGCTCACTGCCGCTTCCAGTTCGGGAACCTGTGTCGA 2885
 Db |||||
 Qy 176 ACTCAGATTAATTCGTTGGCTCACTGCCGCTTCCAGTTCGGGAACCTGTGTCGA 235
 Db |||||
 Qy 2886 GCTGCATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTTTC 2945
 Db |||||
 Qy 236 GCTGCATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTTTC 295
 Db |||||
 Qy 2946 CGCTTCTCTCCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCAGC 3005
 Db |||||
 Qy 296 CGCTTCTCTCCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCAGC 355
 Db |||||
 Qy 3006 TCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATACCGAGAAAGACAT 3065
 Db |||||
 Qy 356 TCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATACCGAGAAAGACAT 415
 Db |||||
 Qy 3066 GTGAGCAAAAGGCGAGCAAAAGCGCAGGAACCGTAAAGGCGGTTGCTGGCGTTT 3125
 Db |||||
 Qy 416 GTGAGCAAAAGGCGAGCAAAAGCGCAGGAACCGTAAAGGCGGTTGCTGGCGTTT 475
 Db |||||
 Qy 3126 CCATAGGCTCGGCGGCTTGAAGCATCAAAAATCAAGCTCAAGTCAAGGTCGCG 3185
 Db |||||
 Qy 476 CCATAGGCTCGGCGGCTTGAAGCATCAAAAATCAAGCTCAAGTCAAGGTCGCG 535
 Db |||||
 Qy 3186 AAAACCCACAGGACTATAAGATACAGAGGCTTTCCTCGGAAGCTCCCTCGTGGCGTC 3245
 Db |||||
 Qy 536 AAAACCCACAGGACTATAAGATACAGAGGCTTTCCTCGGAAGCTCCCTCGTGGCGTC 595
 Db |||||
 Qy 3246 TCCGTGTTCCGACCTCGGCTTACCGGATACCTGTCCGCTTTCCTCGGAGAGGCT 3305
 Db |||||
 Qy 596 TCCGTGTTCCGACCTCGGCTTACCGGATACCTGTCCGCTTTCCTCGGAGAGGCT 655
 Db |||||
 Qy 3306 GGCGCTTTCATAGCTCAGCTGTAGGTATCTCAGTTCGCTGAGGTGCTGCTCCAA 3365
 Db |||||
 Qy 656 GGCGCTTTCATAGCTCAGCTGTAGGTATCTCAGTTCGCTGAGGTGCTGCTCCAA 715
 Db |||||

Qy 3366 GCTGGGCTGTGTGACGAACCCCGGTTGAGCCGCGCTGCGCTTATCCGGTAACCTA 3425
 Db |||||
 Qy 716 GCTGGGCTGTGTGACGAACCCCGGTTGAGCCGCGCTTATCCGGTAACCTA 775
 Db |||||
 Qy 3426 TCGTCTTGTAGTCCAAACCCCGTAAAGACACGACTTATCGGCACCTGGCAGCAGCCTGGTAA 3485
 Db |||||
 Qy 776 TCGTCTTGTAGTCCAAACCCCGTAAAGACACGACTTATCGGCACCTGGCAGCAGCCTGGTAA 835
 Db |||||
 Qy 3486 CAGGATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGGCTTAA 3545
 Db |||||
 Qy 836 CAGGATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGGCTTAA 895
 Db |||||
 Qy 3546 CTACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 3605
 Db |||||
 Qy 896 CTACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 955
 Db |||||
 Qy 3606 CGGAAAAGAGTGTGTAGTCTTGTATCCGGCAAAACAAACACCCGCTGTGTAGCGTGGT 3665
 Db |||||
 Qy 956 CGGAAAAGAGTGTGTAGTCTTGTATCCGGCAAAACAAACACCCGCTGTGTAGCGTGGT 1015
 Db |||||
 Qy 3666 TTTTGTGTGACGACGAGATTAACGGCGCAAGAAAAGATCTCAGAGATCCCTTTGAT 3725
 Db |||||
 Qy 1016 TTTTGTGTGACGACGAGATTAACGGCGCAAGAAAAGATCTCAGAGATCCCTTTGAT 1075
 Db |||||
 Qy 3726 CTTTCTTACGGGCTGTGACGCTCAGTGGAAACGAAAACCTCAAGTAAAGGATTTTGGTCTAT 3785
 Db |||||
 Qy 1076 CTTTCTTACGGGCTGTGACGCTCAGTGGAAACGAAAACCTCAAGTAAAGGATTTTGGTCTAT 1093
 Db |||||
 Qy 3786 GAGATTAATCTGCAACCAAGAGGCGCATCTGCTCTCCCACTCTGCGATTCGGGGGCGATG 3845
 Db |||||
 Qy 1094 GATGCGCGATAGCCGCTGCTGCTGTTCTTGGATGCGGAGATTTGCACTCCGCGTAGAA 3905
 Db |||||
 Qy 1094 GATGCGCGATAGCCGCTGCTGCTGTTCTTGGATGCGGAGATTTGCACTCCGCGTAGAA 1093
 Db |||||
 Qy 3906 CTCCGCGAGTCTCCAGGCTCAGGCGAGCTGAACCACTCCGCGAGGGGATCGAGCC 3965
 Db |||||
 Qy 1094 CTCCGCGAGTCTCCAGGCTCAGGCGAGCTGAACCACTCCGCGAGGGGATCGAGCC 1093
 Db |||||
 Qy 3966 GGGGTGGCGCAAGAACTCCAGCAGTAGATCCCGCGCTGGAGGATCATCCAGCCGCGGTC 4025
 Db |||||
 Qy 1094 GGGGTGGCGCAAGAACTCCAGCAGTAGATCCCGCGCTGGAGGATCATCCAGCCGCGGTC 1093
 Db |||||
 Qy 4026 CCGGAAAACGATTCGGAAGCCCAACCTTTCATAGAAAGCGCGGTGGAATCGAAATCTCG 4085
 Db |||||
 Qy 1094 CCGGAAAACGATTCGGAAGCCCAACCTTTCATAGAAAGCGCGGTGGAATCGAAATCTCG 1093
 Db |||||
 Qy 4086 TGATGGCAGGTTGGGCGTCTGTTGGTCTGCTATTTTCGAACCCAGAGTCCCGCTCAGAG 4145
 Db |||||
 Qy 1094 TGATGGCAGGTTGGGCGTCTGTTGGTCTGCTATTTTCGAACCCAGAGTCCCGCTCAGAG 1103
 Db |||||
 Qy 4146 AACTCGTCAAGAAAGCGGATAGAAAGCGATGCGCTGCGAATCGGAGCGCGGATACCGTAA 4205
 Db |||||
 Qy 1104 AACTCGTCAAGAAAGCGGATAGAAAGCGATGCGCTGCGAATCGGAGCGCGGATACCGTAA 1163
 Db |||||
 Qy 4206 AGCAGGAGGAGCGGTACGCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGAGTACC 4265
 Db |||||
 Qy 1164 AGCAGGAGGAGCGGTACGCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGAGTACC 1223
 Db |||||
 Qy 4266 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAA 4325
 Db |||||
 Qy 1224 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAA 1283
 Db |||||
 Qy 4326 AAGCGGCAATTTCCACCATGATATTTCGGCAAGCAGGATCCCATGGTCAACGAGAGA 4385
 Db |||||
 Qy 1284 AAGCGGCAATTTCCACCATGATATTTCGGCAAGCAGGATCCCATGGTCAACGAGAGA 1343
 Db |||||
 Qy 4386 TCCTCGCGCTCGGCGATGCGCGCTTGAAGCTGTGGGCAACAGTTTCGCTGCGCGAGCCCC 4445
 Db |||||
 Qy 1344 TCCTCGCGCTCGGCGATGCGCGCTTGAAGCTGTGGGCAACAGTTTCGCTGCGCGAGCCCC 1403
 Db |||||

QY 4446 TGATGCTCTTCTGTCAGATCATCTGATCGACAAGACCGCTTCCATCCGAGTACGTGCT 4505
 DB 1404 TGATGCTCTTCTGTCAGATCATCTGATCGACAAGACCGCTTCCATCCGAGTACGTGCT 1463
 QY 4506 CCGTCGATGCGATGTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4565
 DB 1464 CCGTCGATGCGATGTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
 QY 4566 AGCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4625
 DB 1524 AGCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
 QY 4626 AGGAGATCTCTGCGCGGACATCTGCGGACATGATGATGATGATGATGATGATGATGATGAT 4685
 DB 1584 AGGAGATCTCTGCGCGGACATCTGCGGACATGATGATGATGATGATGATGATGATGATGAT 1643
 QY 4686 AGCTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4745
 DB 1644 AGCTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1703
 QY 4746 TCGTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4805
 DB 1704 TCGTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1763
 QY 4806 CCTGCGCTGAGCGCGGACACCGCGGACATGATGATGATGATGATGATGATGATGATGATGAT 4865
 DB 1764 CCTGCGCTGAGCGCGGACACCGCGGACATGATGATGATGATGATGATGATGATGATGATGAT 1823
 QY 4866 TCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4925
 DB 1824 TCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883
 QY 4926 TCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4985
 DB 1884 TCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1943
 QY 4986 CAGATCTCTGCGGACGAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5045
 DB 1944 CAGATCTCTGCGGACGAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003
 QY 5046 GAGGCGGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5105
 DB 2004 GAGGCGGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2063
 QY 5106 TATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5146
 DB 2064 AACTGTTGGGAGGCGATGCTGCTGCGGCGCTCTCTGCTATT 2104

RESULT 11

AAV40796

AAV40796 standard; DNA; 3599 BP.

NC AAV40796;

XX 23-SEP-1998 (first entry)

DT Actual sequence of IGF-I containing plasmid pIG0552.

DE IGF-I; insulin-like growth factor I; skeletal alpha-actin gene promoter;

XX muscle atrophy; diabetes; osteoporosis; growth disorder; therapy; AIDS;

KW Chacot-Marie-Tooth disease; atherogenesis; haemophilia; neuropathy; ss.

XX Synthetic.

OS

XX W09824922-AL.

XX 11-JUN-1998.

ED

XX 01-DEC-1997; 97WO-US21852.

XX 19-NOV-1997; 97US-0974572.

XX 02-DEC-1996; 96US-0031539.

PR

XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA (GENE-) GENEMEDICINE INC.
 XX Coleman M, Demayo FJ, Schwartz R;
 PI MPI; 1998-333339/29.
 XX New vector for expression of insulin-like growth factor-I -
 PT containing a skeletal alpha-actin gene promoter, IGF-I coding
 PT sequences and a 3' region from growth hormone 3'-UTR
 XX Claim 16; Page 37-41; 115pp; English.
 XX This sequence is the actual sequence for pIG0552 which is an example of
 a vector of the invention. The vector is for expression of a nucleic acid
 sequence in a cell, and comprises: (a) a nucleic acid cassette containing
 a sequence encoding insulin-like growth factor-I (IGF-I); (b) a
 5' flanking region including one or more sequences necessary for a
 expression of the nucleic acid cassette, including a promoter from a
 skeletal alpha-actin gene; (c) a linker connecting the 5' flanking region
 to a nucleic acid, the linker having a position for inserting the nucleic
 acid cassette, and lacking the coding sequence of a gene with which it is
 naturally associated; and (d) a 3' flanking region, including a
 3' untranslated region or a 3' non coding region or both, where the
 3' flanking region is 3' to the position for inserting the nucleic acid
 cassette and comprises a sequence from a growth hormone 3'-UTR. The
 vector can provide for efficient IGF-I expression, particularly in gene
 therapy. It can be used for the delivery of IGF-I for treating diseases
 such as muscle atrophy, diabetes, neuropathy, osteoporosis, and growth
 disorders. They can be used for treating peripheral neuropathies and
 resulting from diabetes, genetic disease such as Type I or Type II
 diabetes, genetic disease such as Chacot-Marie-Tooth disease, AIDS,
 atherogenesis, atherosclerotic, cardiovascular, cerebrovascular, or
 peripheral vascular disease, haemophilia, inflammation and side-effects
 from anti-cancer and anti-viral drugs. The vectors can also be used to
 create transgenic animals for research or livestock improvement.

XX Sequence 3599 BP; 733 A; 1098 C; 1042 G; 726 T; 0 other;

Query Match 30.4%; Score 1603.8; DB 19; Length 3599;

Best Local Similarity 82.8%; Pred. No. 1.1e-229; Indels 392; Gaps 1;

Matches 2022; Conservative 0; Mismatches 27;

QY 2706 CTAGAGCTAATCATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2765
 DB 56 CTTGGCGTAAATCATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 115
 QY 2766 ACAACATACGAGCGGAGCATAAAGTAAAGCGCTGGGGTGCCTAATGAGTACGAGCTA 2825
 DB 116 ACAACATACGAGCGGAGCATAAAGTAAAGCGCTGGGGTGCCTAATGAGTACGAGCTA 175
 QY 2826 ACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2885
 DB 176 ACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235
 QY 2886 GCTGATTAATGATCGCCAAACGCGCGGAGAGGCGGTTTTCGCTAATGGGCGCTCTTC 2945
 DB 236 GCTGATTAATGATCGCCAAACGCGCGGAGAGGCGGTTTTCGCTAATGGGCGCTCTTC 295
 QY 2946 CGCTTCCTCGCTCAGTACGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
 DB 296 CGCTTCCTCGCTCAGTACGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 355
 QY 3006 TCATCGAGGCGGTAATACGCTTATCCACAGATCAGGGGATTAACCGAGGAACAT 3065
 DB 356 TCATCGAGGCGGTAATACGCTTATCCACAGATCAGGGGATTAACCGAGGAACAT 415
 QY 3066 GTGAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAA 3125
 DB 416 GTGAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAA 475
 QY 3126 CCATAGGCTCCGCCCCCTCGAGGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCG 3185

Db 476 CCATAGCTCCGCCCCCTGAGAGATACAAAAATCGACGCTCAAGTCAGAGGTGGCG 535
Qy 3186 AAACCCGACAGGACTATAAGATACAGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGGCTC 3245
Db 536 AAACCCGACAGGACTATAAGATACAGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGGCTC 595
Qy 3246 TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGGAAGCGT 3305
Db 596 TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGGAAGCGT 655
Qy 3306 GGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAA 3365
Db 656 GGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAA 715
Qy 3366 GCTGGGCTGTGTCACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTA 3425
Db 716 GCTGGGCTGTGTCACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTA 775
Qy 3426 TCCTGTTGAGTCCAAACCCGTAAGACACAGCTTATCGCCACTGCGACGACCACTGGTAA 3485
Db 776 TCCTGTTGAGTCCAAACCCGTAAGACACAGCTTATCGCCACTGCGACGACCACTGGTAA 835
Qy 3486 CAGGATTAGCAGAGCGATGTAGGCGGTGCTACAGATCTTGAAGTGGTGGCTAA 3545
Db 836 CAGGATTAGCAGAGCGATGTAGGCGGTGCTACAGATCTTGAAGTGGTGGCTAA 895
Qy 3546 CTACGGCTACACAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTT 3605
Db 896 CTACGGCTACACAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTT 955
Qy 3606 CGGAAGAAGTGTGAGTCTTGTATCCGGCAACAAACACACCGCTGAGCGGTGGTTT 3665
Db 956 CGGAAGAAGTGTGAGTCTTGTATCCGGCAACAAACACACCGCTGAGCGGTGGTTT 1015
Qy 3666 TTTTGTGTAAGCAGCAGATTAACCGCAGAAAGAAAGATCTCAAGAGATCTTTGAT 3725
Db 1016 TTTTGTGTAAGCAGCAGATTAACCGCAGAAAGAAAGATCTCAAGAGATCTTTGAT 1075
Qy 3726 CTTTCTACGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTAAGGGATTTGGTCAT 3785
Db 1076 CTTTCTACGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTAAGGGATTTGGTCAT 1093
Qy 3786 GAGATTATCTGACCAAGAGCGGCATCTGCGCTCCCACTCTCTGAGTTCGGGGGCATG 3845
Db 1094 ----- 1093
Qy 3846 GATCGCGATAGCGCTGCTGCTTTCCTGATGCGAGCGGATTTGCACTGCGGTAGAA 3905
Db 1094 ----- 1093
Qy 3906 CTCGCGAGTCTGCCAGCTCAGGACGAGCTGAACCACTCGGAGGGGATCGAGCCC 3965
Db 1094 ----- 1093
Qy 3966 GGGTGGGGAAGAACTCCAGCATGAGATCCCGCGCTGAGGATCATCAGCGCGCTC 4025
Db 1094 ----- 1093
Qy 4026 CCGGAAACGATTCCGAAGCCAACTTTATAGAGGCGCGGTGGAAATCGAAATCTCG 4085
Db 1094 ----- 1093
Qy 4086 TGATGGAGGTGGGCGTCTGTTGGTCCGTCAATTTCCGAACCCAGAGTCCCGCTCAGAAG 4145
Db 1094 ----- CGCTCAGAAG 1103
Qy 4146 AACTCGTCAAGAGGCGATAGAGCGATGCGCTCGAATCGGAGCGGCGATACCGTAA 4205
Db 1104 AACTCGTCAAGAGGCGATAGAGCGATGCGCTCGAATCGGAGCGGCGATACCGTAA 1163
Qy 4206 AGCAGAGGAGCGGTACGCCATTCGCGGCTCTCAGCAATATCACGGGTAGCC 4265

Db 1164 AGCAGAGGAGCGGTACGCCCATTTGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC 1223
Qy 4266 AAGCTATATCTCTGATAGCGTTCGGCCACACCCAGCGCGGCACAGTGCATGAATCAGAA 4325
Db 1224 AAGCTATATCTCTGATAGCGTTCGGCCACACCCAGCGCGGCACAGTGCATGAATCAGAA 1283
Qy 4326 AAGCGGCATTTTCCACCATGATATTGGCAAGCAGCAGCATCGCCATGGTCAAGAGAGA 4385
Db 1284 AAGCGGCATTTTCCACCATGATATTGGCAAGCAGCAGCATCGCCATGGTCAAGAGAGA 1343
Qy 4386 TCCTGCGCTCGGCGATCGCGGCTTGAAGCTTGGCAACAGTTCGGTGGCGGAGGCCCC 4445
Db 1344 TCCTGCGCTCGGCGATCGCGGCTTGAAGCTTGGCAACAGTTCGGTGGCGGAGGCCCC 1403
Qy 4446 TGATGCTCTTCGTCAGATCATCTGATCGAACAAGCGGCTTCCATTCGAGTACGTCT 4505
Db 1404 TGATGCTCTTCGTCAGATCATCTGATCGAACAAGCGGCTTCCATTCGAGTACGTCT 1463
Qy 4506 CGCTCGATCGGATGTTTCGCTTGGTGGTGAATCGGCAAGTTCAGCGATCAAGCGTATGC 4565
Db 1464 CGCTCGATCGGATGTTTCGCTTGGTGGTGAATCGGCAAGTTCAGCGATCAAGCGTATGC 1523
Qy 4566 AGCGCGCGCATTTGCATCAGCCATGATGATATCTTCTCGGAGAGCAAGGTGAGATGAC 4625
Db 1524 AGCGCGCGCATTTGCATCAGCCATGATGATATCTTCTCGGAGAGCAAGGTGAGATGAC 1583
Qy 4626 AGGAGATCTCTGCGCGCACTTTCGCCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACA 4685
Db 1584 AGGAGATCTCTGCGCGCACTTTCGCCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACA 1643
Qy 4686 AGCTCAGCAGCAGTTCGCGCAAGGAGCAAGCGCTGCGTGGCGCAGCAGATAGCGCGTGC 4745
Db 1644 AGCTCAGCAGCAGTTCGCGCAAGGAGCAAGCGCTGCGTGGCGCAGCAGATAGCGCGTGC 1703
Qy 4746 TCCTCTGAGTTCATTCAGGCGCACCGGACAGTTCGCTTGAACAAAGAAACCGGCGCG 1763
Db 1704 TCCTCTGAGTTCATTCAGGCGCACCGGACAGTTCGCTTGAACAAAGAAACCGGCGCG 1763
Qy 4806 CCTGCGCTGACAGCGGAAACACGCGGCAATCAGCAGCGGATTTGCTGTTGCGCCAG 4865
Db 1764 CCTGCGCTGACAGCGGAAACACGCGGCAATCAGCAGCGGATTTGCTGTTGCGCCAG 1823
Qy 4866 TCATAGCGAATAGCTCTCCACCGAGCGCGGAGAACTGGTGCAATCCATCTTGT 4925
Db 1824 TCATAGCGAATAGCTCTCCACCGAGCGCGGAGAACTGGTGCAATCCATCTTGT 1883
Qy 4926 TCAATCATCGGAAACGATCTCTCATCTCTTGTATCAGATCTTATCCCTGCGCCAT 4985
Db 1884 TCAATCATCGGAAACGATCTCTCATCTCTTGTATCAGATCTTATCCCTGCGCCAT 1943
Qy 4986 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCCAACCTTACCA 5045
Db 1944 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCCAACCTTACCA 2003
Qy 5046 GAGGGCGCGGAGCTGGCAATTCGGTTCGTTGCTGCTCCTATTAACCGGCGGCTAGC 5105
Db 2004 GAGGGCGCGGAGCTGGCAATTCGGTTCGTTGCTGCTCCTATTAACCGGCGGCTAGC 2063
Qy 5106 TATCGGCATGTAAGCCCACTGCAAGCTACCTGCTTCTTCTT 5146
Db 2064 AACTGTTGGGAGGCGATCGGTGGCGGCTCTTCGCTATT 2104

RESULT 12
AAV50427
ID AAV50427 standard; DNA; 3600 BP.
XX
AC AAV50427;
XX
DT 26-OCT-1998 (first entry)
XX
DB Plasmid pIG0552 upper expected sequence nucleotides 1-3600.
XX

KW Human; IGF-1; insulin-like growth factor 1; urinary incontinence;
 KW gene therapy; neurotrophic factor; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX WO9833529-A1.
 XX 06-AUG-1998.
 XX 04-FEB-1998; 98WO-US02051.
 XX 04-FEB-1997; 97US-0036862.
 XX (GENE-) GENEMEDICINE INC.
 XX Coleman M;
 XX WPI; 1998-437184/37.
 XX Treatment of urinary incontinence - by delivering nucleic acid
 XX vector for expression of growth factor or neurotrophic factor in
 XX tissue(s)
 XX Disclosure; Page 103-105; 117pp; English.
 XX A method has been developed of treating urinary incontinence (UI) in
 XX mammals. The method comprises delivering a nucleic acid vector for the
 XX expression of a growth factor or neurotrophic factor in a tissue or
 XX tissues. The present sequence represents the expected upper sequence
 XX of plasmid pIG0552 nucleotides 1-3600 from the present invention. Due
 XX to the growth and stimulatory effects of growth factors and neurotrophic
 XX factors, introducing these factors to degenerated muscles in the
 XX urinary system can improve UI by enhancing both their integrity and
 XX neural innervation.
 XX Sequence 3600 BP; 733 A; 1099 C; 1042 G; 726 T; 0 other;
 XX
 XX Query Match 30.3%; Score 1602.2; DB 19; Length 3600;
 XX Best Local Similarity 82.8%; Pred. No. 1.9e-239;
 XX Matches 2021; Conservative 0; Mismatches 28; Indels 392; Gaps 1;
 XX
 XX 2706 CTAGAGCTAATGCTAGCTAGCTTTCTTCTGTGTAATGTTATCCGCTCACAATCC 2765
 XX 57 CTTCGGCTAATCATGTCATAGCTTTCTTCTGTGTAATGTTATCCGCTCACAATCC 116
 XX 2766 ACAACAATACGAGCGGAGCATAAAGTGAAGCTTGGGCTGCCTTAATGAGTGAGCTA 2825
 XX 117 ACAACAATACGAGCGGAGCATAAAGTGAAGCTTGGGCTGCCTTAATGAGTGAGCTA 176
 XX 2826 ACTCACAATTAATGGCTTGGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCCA 2885
 XX 177 ACTCACAATTAATGGCTTGGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCCA 236
 XX 2886 GCTGCATTAATGAATCGGCAACGCGGGAGAGGCGGTTTGGGTAATGGGCGCTCTTC 2945
 XX 237 GCTGCATTAATGAATCGGCAACGCGGGAGAGGCGGTTTGGGTAATGGGCGCTCTTC 296
 XX 2946 CGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
 XX 297 CGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 356
 XX 3006 TCACTCAAAAGCGGTAATACGCTTATCCACAGATCAGGGGTAACGCGAGGAAGACAT 3065
 XX 357 TCACTCAAAAGCGGTAATACGCTTATCCACAGATCAGGGGTAACGCGAGGAAGACAT 416
 XX 3066 GTGAGCAAAAGCGGCAAAAGCGGAGAACCGGTAAAGGCGCGGCTTGTGCGGTTTTT 3125
 XX 417 GTGAGCAAAAGCGGCAAAAGCGGAGAACCGGTAAAGGCGCGGCTTGTGCGGTTTTT 476
 XX 3126 CCAATAGGCTCGGCGCGCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGGTGGCG 3185
 XX 477 CCAATAGGCTCGGCGCGCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGGTGGCG 536

QY 3186 AAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGTCGCTC 3245
 DB 537 AAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTGTCGCTC 596
 QY 3246 TCCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGCGT 3305
 DB 597 TCCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGCGT 656
 QY 3306 GCGGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTCCAA 3365
 DB 657 GCGGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTCCAA 716
 QY 3366 GCTGGGCTGTGTGACAGCAACCCCGCTTACGCGCGAGCGCTGCGCTTATCCGGAACCTA 3425
 DB 717 GCTGGGCTGTGTGACAGCAACCCCGCTTACGCGCGAGCGCTGCGCTTATCCGGAACCTA 776
 QY 3426 TCGTCTTCAGTCCCAACCCGCTAAGACAGCACTTATCCCACTGGCAGCAGCCTGTTAA 3485
 DB 777 TCGTCTTCAGTCCCAACCCGCTAAGACAGCACTTATCCCACTGGCAGCAGCCTGTTAA 836
 QY 3486 CAGGATTAAGCAGAGGAGTATGAGCGGTGCTACAGAGTTCTTGAAGTGTGCGCTAA 3545
 DB 837 CAGGATTAAGCAGAGGAGTATGAGCGGTGCTACAGAGTTCTTGAAGTGTGCGCTAA 896
 QY 3546 CTACGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTT 3605
 DB 897 CTACGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTT 956
 QY 3606 CGGAAAAGAGTTGGTGTAGCTTGTATCCGCAACCAACCAACCGCTGTGAGCGGTGTT 3665
 DB 957 CGGAAAAGAGTTGGTGTAGCTTGTATCCGCAACCAACCAACCGCTGTGAGCGGTGTT 1016
 QY 3666 TTTGTTTGAAGCAGCAGATTAACCGCAGCAAAAAAGGATCTCAAGAGATCTTTGAT 3725
 DB 1017 TTTGTTTGAAGCAGCAGATTAACCGCAGCAAAAAAGGATCTCAAGAGATCTTTGAT 1076
 QY 3726 CTTTCTACGGGCTCTGACGCTCAGTGAACGAAACTCAGCTTAAGGATTTTGGTCAT 3785
 DB 1077 CTTTCTACGGGCTCTGACGCTCAGTGAACGAAACTCAGCTTAAGGATTTTGGTCAT 1094
 QY 3786 GAGATTAATCGTGACCAAAAGCGGCATCTGCTCCCTCTGCTGAGTTCCGGGCGATG 3845
 DB 1095 ----- 1094
 QY 3846 GATGCGCGGATAGCGCTGCTGTTTCTGGATGCCGACGGATTTGCACTGCCGCTAGAA 3905
 DB 1095 ----- 1094
 QY 3906 CTCGCGAGGTCGTCAGCCTCAGGACAGCTGAACCAACTCGCGAGGGATCGAGCC 3965
 DB 1095 ----- 1094
 QY 3966 GGGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGAGGATCATCCAGCCGCGCTC 4025
 DB 1095 ----- 1094
 QY 4026 CGGAAAACGATTCGAAGCCCAACCTTTTATAGAAGCGCGGTGGAATCGAATCTCG 4085
 DB 1095 ----- 1094
 QY 4086 TGATGGCAGGTTGGGCTGCTGTTGGTGGTTCATTTCGAACCCAGAGTCCCTCAGAG 4145
 DB 1095 ----- CCGTCAAGAG 1104
 QY 4146 AACTCGTCAAGAGCGGATAGAGCGGATGCGGTGCGAATCGGAGCGCGGATACCGTAA 4205
 DB 1105 AACTCGTCAAGAGCGGATAGAGCGGATGCGGTGCGAATCGGAGCGCGGATACCGTAA 1164
 QY 4206 AGCAGAGAGAGCGGTGAGCGGCTTCCCGCCAAAGCTCTTTCAGCAATATCAAGGTTAGCC 4265
 DB 1165 AGCAGAGAGCGGTGAGCGGCTTCCCGCCAAAGCTCTTTCAGCAATATCAAGGTTAGCC 1224

IGF-I; insulin-like growth factor I; skeletal alpha-actin gene promoter; muscle atrophy; diabetes; osteoporosis; growth disorder; therapy; AIDS;

RESULT 14

AXX88055
ID AAX88055 standard; DNA: 5707 BP.

AC AAX88055:

DT 08-SEP-1999 (first entry)

Plasmid pIG0335 DNA.

Plasmid pIG0335; expression vector; treatment; disease;

KW RNA stability element; gene therapy; muscle atrophy; neurological; RNA stability element; gene therapy; muscle atrophy; neurological;

muscular disease; systemic disease; aging; trophic factor; haemophilia;

clotting factor; atherogenesis; atherosclerotic; cardiovascular; KW

KW cerebrovascular; peripheral-vascular disease; hormone deficiency;
KW diabetes: transgenic animal: carcinogen: regulatory element:

KW diabetes; transgenic animal; carcinogen; regulatory element;
KW livestock improvement; immune response; ds.

XX
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707

OS Synthetic.

PN US5925564-A.

XX	20	100
XX	20	100

PD 20-JUL-1999.
XX

XX
PF 07-JUN-1995: 95JIS-0472809

PF 07-JUN-1995; 95US-04/2809.
XX

PR 07-JUN-1995; 95US-0472809.

06-NOV-1991; 91US-0789919.

PR 09-MAR-1994; 94US-0209846.

XX

PA (BAYU) BAYLOR COLLEGE MEDICINE.
yy

XX
PI DeMAYO F.T O'Malley BW Schwartz RJ.

XX	Demayo RJ, O'Malley BW, Schwartz RJ;
XX	WPI: 1999-418276/35.
XX	PT New expression vector system useful for gene therapy
XX	PT Disclosure; Fig 17-17P; 67pp; English.
XX	This invention describes novel expression vector systems containing RNA
CC	stability elements from 3' flanking sequences used for establishing
CC	expression of a nucleic acid sequence within a tissue. The vectors also
CC	facilitate enhanced expression in tissues and target expression with
CC	tissue specificity. The expression vectors can be used to treat diseases
CC	through gene therapy by targeting the vector to specific tissues.

Diagnosis and therapy by targeting the vector to specific tissues. Diseases that can be treated include muscle atrophy associated with neurological, muscular or systemic disease, aging by causing tissues to express trophic factors, haemophilia by causing tissues to express and secrete clotting factor into the circulation, atherosclerosis and arterosclerotic cardiovascular, cerebrovascular or peripheral-vascular disease by causing tissues to express factors involved in tissue metabolism. They can be used to replace genes of inherited genetic defects or acquired hormone deficiencies e.g. diabetes. To transform cells to produce particular proteins or RNA in vitro. To create transgenic animals which can be used for research into human diseases, assessing novel therapeutic methods, assessing the effect of chemical and physical carcinogens and for studying the effect of genes and genetic regulatory elements or livestock improvement. They can be used to induce an immune response. These vectors provide controlled expression of the genes they carry and produce a significantly high level of expression. Using 3'UTR sequences reduces the decay rates of the mRNAs encoded by the vectors which causes increased expression.

[illegible]

SQ Sequence 5707 BP; 1221 A; 1634 C; 1576 G; 1276 T; 0 other;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match 30.3%; Score 1602.2; DB 20; Length 5707;

Best Local Similarity	82.8%;	Pred. No. 1.9e-239;
Matches 2021: Conservative	0. Mismatched	38.
Matches 2021: Indol	38.	38.

Matches 2021; Conservative 0; Mismatches 28; Indels 392; Gaps 1;

QV 2706 CTAGACGTAATCATGGTCATAGCTGTTCCCTGTGTGAAATTGTTATCCGCTCACAATTCC 2765

2700 CAGACCGGATCATGGATCATAGACCGTATCCCTGTGTGAAATGTATATCCGATCAGAAATCC 2760

100

3514	DB	CTTGGCGTAAATCATGCTCATAGCTGTTTCTGTGTGAATAATTTGTTATTCGCGCTCACAATTC	3571
2766	QY	ACACAAATATACGACCGGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTACGCTA	2825
3574	DB	ACACAAATATACGACCGGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTACGCTA	3633
2826	QY	ACTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTTCAGTTCGGGAAACCTGTGCTGCCA	2885
3634	DB	ACTCACATTAATTTGCGTTGCGCTCACTGCCCGCTTTTCAGTTCGGGAAACCTGTGCTGCCA	3693
2886	QY	GCTGCATTAATGAATCGGGCAACCGCGCGGGAGAGCGGTTTGGGTATTTGGCGGCTCTTC	2945
3694	DB	GCTGCATTAATGAATCGGGCAACCGCGCGGGAGAGCGGTTTGGGTATTTGGCGGCTCTTC	3753
2946	QY	CGCTTCTCCGCTCACTGACTCTCGTGGCGCTCGGTTCGCTCGCGGCGAGCGGTATCAGC	3005
3754	DB	CGCTTCTCCGCTCACTGACTCTCGTGGCGCTCGGTTCGCTCGCGGCGAGGTATCAGC	3813
3006	QY	TCACTCAAAAGCGGTAAATACGGTTATTCACAGAAATCAGGGATTAACCGAGGAAGAACAT	3065
3814	DB	TCACTCAAAAGCGGTAAATACGGTTATTCACAGAAATCAGGGATTAACCGAGGAAGAACAT	3873
3066	QY	GTGAGCAAAAGGCCAGCAAAAGCCAGGAAACCGTAAAAAGCCCGTGTGCGCGTTTTT	3125
3874	DB	GTGAGCAAAAGGCCAGCAAAAGCCAGGAAACCGTAAAAAGCCCGTGTGCGCGTTTTT	3933
3126	QY	CCATAGGCTCGGCCCTTGCAGCAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG	3185
3934	DB	CCATAGGCTCGGCCCTTGCAGCAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG	3993
3186	QY	AAACCCACAGGACTAATAGATACACAGCGTTTCCGCCCTCGAAGCTCCCTCGTGGCGCTC	3245
3994	DB	AAACCCACAGGACTAATAGATACACAGCGTTTCCGCCCTCGAAGCTCCCTCGTGGCGCTC	4053
3246	QY	TCCTGTTCCGACCCGCGGTACCGGATACTGTCGCGCTTTCTCCGTCGGGAACGCT	3305
4054	DB	TCCTGTTCCGACCCGCGGTACCGGATACTGTCGCGCTTTCTCCGTCGGGAACGCT	4113
3306	QY	GGCGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTTCGGTGTAGTCTGTCGCTCCAA	3365
4114	DB	GGCGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTTCGGTGTAGTCTGTCGCTCCAA	4173
3366	QY	GCTGGGCTGTGTCAGCAAAACCCCGGTTTCAGGCCGACCGCTGCGCCTTATCCGGTAACTA	3425
4174	DB	GCTGGGCTGTGTCAGCAAAACCCCGGTTTCAGGCCGACCGCTGCGCCTTATCCGGTAACTA	4233
3426	QY	TCGCTTTGAGTCCAAACCGGTAAAGCACAGCTATTCGCCACTGCGACGACGCACTGGTAA	3485
4234	DB	TCGCTTTGAGTCCAAACCGGTAAAGCACAGCTATTCGCCACTGCGACGACGCACTGGTAA	4293
3486	QY	CAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCGCTAA	3545
4294	DB	CAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCGCTAA	4353
3546	QY	CTACGGGTACACTAGAGAAGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTTACCTT	3605
4354	DB	CTACGGGTACACTAGAGAAGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTTACCTT	4413
3606	QY	CGGAAAAGAGGTTGGTAGCTCTTGATTCGGGCAAAACAAACCAACCGCTGGTAGCGGTGTT	3665
4414	DB	CGGAAAAGAGGTTGGTAGCTCTTGATTCGGGCAAAACAAACCAACCGCTGGTAGCGGTGTT	4473
3666	QY	TTTTGTTTGCAGCAGCAGATTTACCGCGAGAAAAGAGATCTCAAGAGATCCCTTTGAT	3725
4474	DB	TTTTGTTTGCAGCAGCAGATTTACCGCGAGAAAAGAGATCTCAAGAGATCCCTTTGAT	4533
3726	QY	CTTTTCTACGGGGTCTGAAGCTCAGCTGGAACGAAACCTCAGTTAAAGGATTTTGGTTCAT	3785
4534	DB	CTTTTCTACGGGGTCTGAAGCTCAGCTGGAACGAAACCTCAGTTAAAGGATTTTGGTTCAT	4551
3786	QY	GAGATTTATCGTCCACCAAAAGCGGCCATCGTGCTCTCCCACTCTCTGCAAGTTCCGGGGCATG	3845

Db 120 ACACACATACGACCGGAGCATAAAGTGTAAAGCCTGGGTGCTTAATGAGTACGCTA 179
Qy 2826 ACTCACATTAATTTGGGTGGCTCACTGCCCGGTTTTCAGTCCGGAAACCTGTGTGCTCA 2885
Db 180 ACTCACATTAATTTGGGTGGCTCACTGCCCGGTTTTCAGTCCGGAAACCTGTGTGCTCA 239
Qy 2886 GCTGCAATTAATTAATCGGCAACCGCGGGGAGAGCGGTTTTCGCTATTGGGCGCTTTC 2945
Db 240 GCTGCAATTAATTAATCGGCAACCGCGGGGAGAGCGGTTTTCGCTATTGGGCGCTTTC 299
Qy 2946 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
Db 300 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 359
Qy 3006 TCACTCAAGCGGCTAATCGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACAT 3065
Db 360 TCACTCAAGCGGCTAATCGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACAT 419
Qy 3066 GTGAGCAAAAGCGCAGCAAAAGCGCAGCAACCGTAAAGCGCGGTTGCTGGGCTTTT 3125
Db 420 GTGAGCAAAAGCGCAGCAAAAGCGCAGCAACCGTAAAGCGCGGTTGCTGGGCTTTT 479
Qy 3126 CCATAGCTCCGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCG 3185
Db 480 CCATAGCTCCGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCG 539
Qy 3186 ABAACCGCAGCAAGTAAAGATACAGCGGTTTCCCGCTGGAAGTCCCTCGTGGCTC 3245
Db 540 ABAACCGCAGCAAGTAAAGATACAGCGGTTTCCCGCTGGAAGTCCCTCGTGGCTC 599
Qy 3246 TCCTGTTCCGACCCCTCGCTTACCGGATACGCTGTCGCGCTTTCCTCCCTCGGGAAGCGT 3305
Db 600 TCCTGTTCCGACCCCTCGCTTACCGGATACGCTGTCGCGCTTTCCTCCCTCGGGAAGCGT 659
Qy 3306 GCGCTTTCATAGCTCAAGTGTAGTATCTAGTTCGCTGAGTGTCTGCTGCTCA 3365
Db 660 GCGCTTTCATAGCTCAAGTGTAGTATCTAGTTCGCTGAGTGTCTGCTGCTCA 719
Qy 3366 GCTGGGCTGTGTGACGAACCCCGCTTACGCGCCGCTGCGCTTATCCGCTA 3425
Db 720 GCTGGGCTGTGTGACGAACCCCGCTTACGCGCCGCTGCGCTTATCCGCTA 779
Qy 3426 TCCTGTTGACCAACCGGTAGACACACTTATCGCACTGCGCAGCAGCACTGTGTA 3485
Db 780 TCCTGTTGACCAACCGGTAGACACACTTATCGCACTGCGCAGCAGCACTGTGTA 839
Qy 3486 CAGGATAGCAGGAGGTATGTAGCGGCTGTACAGAGTTCTTTGAAGTGTGCGCTTAA 3545
Db 840 CAGGATAGCAGGAGGTATGTAGCGGCTGTACAGAGTTCTTTGAAGTGTGCGCTTAA 899
Qy 3546 CTAGGCTACACTAGAGACAGTATTTGGTATCTGCTGCTGCTGAGCCAGTTACCTT 3605
Db 900 CTAGGCTACACTAGAGACAGTATTTGGTATCTGCTGCTGCTGAGCCAGTTACCTT 959
Qy 3606 CGGAAAGAGTGTGTAGTCTTTGATCCGGGCAACCAACCCGCTGTAGCGGTGTTT 3665
Db 960 CGGAAAGAGTGTGTAGTCTTTGATCCGGGCAACCAACCCGCTGTAGCGGTGTTT 1019
Qy 3666 TTTTGTTCAGCAGCAGTATACGGCAGAAAGAGTCTCAAGAGTCTTTGAT 3725
Db 1020 TTTTGTTCAGCAGCAGTATACGGCAGAAAGAGTCTCAAGAGTCTTTGAT 1079
Qy 3726 CTTTCTACCGGCTGTGAGCTCAGTGGAAACGAAACCTCAGTTAAGGATTTTGTGAT 3785
Db 1080 CTTTCTACCGGCTGTGAGCTCAGTGGAAACGAAACCTCAGTTAAGGATTTTGTGAT 1097
Qy 3786 GAGATTATCTGACCAAGAGCGCCATCGTGCTCCCGACTCTCTGCTGCGGCGCATG 3845
Db 1098 ----- 1097
Qy 3846 GATGCGCGGATAGCCGCTGCTGTTTCTGAGTGGCAGCGGATTTGCACTGCCGATGAA 3905
Db 1098 ----- 1097

Qy 3906 CTTCCGCGAGGTCGTCCAGCCTCAGGCAAGAGTGAACCAACTCGGAGGGGATCGAGCCC 3965
Db 1098 ----- 1097
Qy 3966 GGGGTGGGCAAGAACTCCAGCATGAGATCCCGGCTGGAGGATCATCCAGCCGGCTC 4025
Db 1098 ----- 1097
Qy 4026 CCGGAAACGATTCGGAAGCCCAACCTTTTATAGAAGCGCGGTGGATCGAAATCTCG 4085
Db 1098 ----- 1097
Qy 4086 TGAATGCGAGTTGGGCGTGGCTTGGTGGTCAATTCGAACCCAGAGTCCCGCTCAGAAG 4145
Db 1098 -----CGCTCAGAAG 1107
Qy 4146 AACTGCTCAGAAAGCGATAGAAAGCGATGCGTTCGMAATCGGAGCGGGGATACCGTAA 4205
Db 1108 AACTGCTCAGAAAGCGATAGAAAGCGATGCGTTCGMAATCGGAGCGGGGATACCGTAA 1167
Qy 4206 AGCAGAGGAAGCGGTCAAGCCATTTGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC 4265
Db 1168 AGCAGAGGAAGCGGTCAAGCCATTTGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC 1227
Qy 4266 AAGCGGCTATTTCCACCATGATATTCGGCAAGCAGGATCGCCATCGCTCAGCAGGA 4325
Db 1288 AAGCGGCTATTTCCACCATGATATTCGGCAAGCAGGATCGCCATCGCTCAGCAGGA 1347
Qy 4386 TCCTCGCGTGGGCTGCGGCTGAGCTGCGCAAGCTGCGGCTGCGGCTGCGGCTGCGGCT 4445
Db 1348 TCCTCGCGTGGGCTGCGGCTGAGCTGCGCAAGCTGCGGCTGCGGCTGCGGCTGCGGCT 1407
Qy 4446 TGATGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 4505
Db 1408 TGATGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 1467
Qy 4506 CGCTCGATCGGATGTTTCGCTTGGTTCGAAATCGGAGGAGTCCGATCAAGCGTATGC 4565
Db 1468 CGCTCGATCGGATGTTTCGCTTGGTTCGAAATCGGAGGAGTCCGATCAAGCGTATGC 1527
Qy 4566 AGCCCGCAGATTCGATCAGCCATGATGATATTTCTCGGAGGAGCAAGGTGAGATGAC 4625
Db 1528 AGCCCGCAGATTCGATCAGCCATGATGATATTTCTCGGAGGAGCAAGGTGAGATGAC 1587
Qy 4626 AGGAGATCTGCGCGGCTTCCGCTTGGTTCGAAATCGGAGGAGTCCGATCAAGCGTATGC 4685
Db 1588 AGGAGATCTGCGCGGCTTCCGCTTGGTTCGAAATCGGAGGAGTCCGATCAAGCGTATGC 1647
Qy 4686 AGCTCGAGCAGGTGCGCAAGGAACCGCTGTCGCGCAGCAGATAGCCGCTGCTGCT 4745
Db 1648 AGCTCGAGCAGGTGCGCAAGGAACCGCTGTCGCGCAGCAGATAGCCGCTGCTGCT 1707
Qy 4746 TCGTCTCGCTTCAATTCAGGGGCTCCGAGAGTCTGCTTGAACAAAGAAACCGGCGC 4805
Db 1708 TCGTCTCGCTTCAATTCAGGGGCTCCGAGAGTCTGCTTGAACAAAGAAACCGGCGC 1767
Qy 4806 CCGTCTGCTGAGCGCGGAAACCGCGGCTCAGAGAGCGGATGCTGCTGCTGCTGCTGCT 4865
Db 1768 CCGTCTGCTGAGCGCGGAAACCGCGGCTCAGAGAGCGGATGCTGCTGCTGCTGCTGCT 1827
Qy 4866 TCATAGCGGATAGCTCTCCACCCAGCGCGGAGAACTGCTGCTGCTGCTGCTGCTGCT 4925
Db 1828 TCATAGCGGATAGCTCTCCACCCAGCGCGGAGAACTGCTGCTGCTGCTGCTGCTGCT 1887
Qy 4926 TCATAGCGGAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4985
Db 1888 TCATAGCGGAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1947

Tue Feb 17 08:36:51 2004

us-09-921-143-36.rng

QY	4986	CAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGACGGCTTCCCAACCTTACCA	5045
Db	1948	CAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGACGGCTTCCCAACCTTACCA	2007
QY	5046	GAGGGCGCCCCAGCTGGCAATTCCGGTTGCTGTCCATAAAACGGCCAGTCTAGC	5105
Db	2008	GAGGGCGCCCCAGCTGGCAATTCCGGTTGCTGTCCATAAAACGGCCAGTCTAGC	2067
QY	5106	TA	5107
Db	2068	AA	2069

Search completed: February 15, 2004, 11:11:36
Job time : 1314 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 01:53:42 ; Search time 272 Seconds
(without alignments)
8572.894 Million cell updates/sec

Title: US-09-921-143-36
Perfect score: 5283
Sequence: 1 aagttgacattatgcgact.....tgagtgcttcgcgcagcgtg 5283

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 563978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/FACTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/FACTUS_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2240	42.4	8349	4	US-09-186-002-16
2	1793.4	33.9	4800	4	US-09-554-928-1
3	1652.4	31.3	5594	4	US-09-380-190A-29
4	1618.2	30.6	6561	4	US-09-380-190A-30
5	1602.2	30.3	5707	2	US-08-472-809B-8
6	1571.6	29.7	4665	3	US-08-948-378A-7
7	1571.6	29.7	4665	3	US-09-169-425C-7
8	1571.6	29.7	4665	4	US-09-759-960-7
9	1570	29.7	4518	4	US-09-380-190A-26
10	1570	29.7	4886	4	US-09-533-220A-4
11	1570	29.7	6139	2	US-08-751-767A-7
12	1569.8	29.7	8797	2	US-08-723-306-6
13	1569.8	29.7	8797	5	PCT-US96-10041-6
14	1569.8	29.7	11093	2	US-08-723-306-5
15	1569.8	29.7	11093	5	PCT-US96-10041-5
16	1501.4	28.4	6795	4	US-09-380-190A-22
17	1501.4	28.4	8574	4	US-09-554-572-3
18	1501.4	28.4	9093	4	US-09-380-190A-23
19	1501.4	28.4	9093	4	US-09-554-572-2
20	1501.4	28.4	9145	4	US-09-554-572-1
21	1501	28.4	4958	4	US-09-380-190A-20
22	1501	28.4	5724	4	US-09-380-190A-18
23	1501	28.4	5724	4	US-09-380-190A-19
24	1501	28.4	6321	4	US-09-380-190A-17
25	1501	28.4	6321	4	US-09-380-190A-16
26	1501	28.4	6359	4	US-09-380-190A-15
27	1501	28.4	6891	4	US-09-380-190A-16

C	28	1501	28.4	8657	4	US-09-380-190A-14	Sequence 14, Appl
	29	1490.4	28.2	3974	3	US-09-026-343-33	Sequence 33, Appl
	30	1490.4	28.2	3974	3	US-09-042-105-16	Sequence 16, Appl
	31	1490.4	28.2	3974	3	US-09-044-856A-7	Sequence 7, Appl
	32	1490.4	28.2	3974	3	US-09-023-082A-147	Sequence 147, Appl
	33	1490.4	28.2	3974	3	US-09-044-855A-7	Sequence 7, Appl
	34	1490.4	28.2	3974	3	US-09-078-670-4	Sequence 4, Appl
	35	1490.4	28.2	3974	4	US-09-026-408-14	Sequence 14, Appl
	36	1490.4	28.2	3974	4	US-09-362-871-33	Sequence 33, Appl
	37	1490.4	28.2	3974	4	US-09-627-154-4	Sequence 4, Appl
	38	1490.4	28.2	3974	4	US-09-027-287-50	Sequence 50, Appl
	39	1490.4	28.2	3974	4	US-09-437-602-4	Sequence 4, Appl
	40	1490.4	28.2	3974	4	US-09-252-656B-50	Sequence 50, Appl
	41	1490.4	28.2	3984	3	US-09-044-796A-10	Sequence 10, Appl
	42	1490.4	28.2	3984	4	US-09-725-460A-10	Sequence 10, Appl
	43	1490.4	28.2	4208	3	US-09-044-796A-9	Sequence 9, Appl
	44	1490.4	28.2	4208	4	US-09-725-460A-9	Sequence 9, Appl
	45	1490.4	28.2	4208	4	US-09-689-693-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1

US-09-186-002-16
; Sequence 16, Application US/09186002B
; Patent No. 6489542
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Method for Transforming Plants to Express
; TITLE OF INVENTION: delta-Endotoxins
; FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002
; CURRENT APPLICATION NUMBER: US/09/186,002B
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 8349
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: gene
; LOCATION: (3666)..(5573)
; OTHER INFORMATION: completely synthesized
US-09-186-002-16

Query Match	42.4%	Score 2240;	DB 4;	Length 8349;
Best Local Similarity	93.4%	Pred. No. 0;		
Matches 2409;	Conservative 0;	Mismatches 25;	Indels 144;	Gaps 1;
Qy	2706	CTAGACGTAAATCATGGTCAATAGCTGTTTCTCTGTGAAATTTGTTATCCGTCACAAATTC	2765	
Db	5911	CTTGGCGTAATCATGGTCAATAGCTGTTTCTCTGTGAAATTTGTTATCCGTCACAAATTC	5970	
Qy	2766	ACCAACATACGACGCGGAGCATAAGCTAAAGCTGGGTCCTTAATGATGAGCTA	2825	
Db	5971	ACCAACATACGACGCGGAGCATAAGCTAAAGCTGGGTCCTTAATGATGAGCTA	6030	
Qy	2826	ACTCACATTAATTCGTTGGCTCACTGCGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA	2885	
Db	6031	ACTCACATTAATTCGTTGGCTCACTGCGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA	6090	
Qy	2886	GCTGCATTAATGATCGGCCCAACGCGCGGAGAGCGGTTTCGTTATTCGGGCTCTTC	2945	
Db	6091	GCTGCATTAATGATCGGCCCAACGCGCGGAGAGCGGTTTCGTTATTCGGGCTCTTC	6150	
Qy	2946	CGCTTCCTCGCTCACTCACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGC	3005	
Db	6151	CGCTTCCTCGCTCACTCACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGC	6210	
Qy	3006	TCATCTAAGCGCGGTAATACGGTTATCCACAGATCAGGCGGATACCGCAGGAAGACAT	3065	

Db 6211 TCACTCAAAGCGCGTAAATACGGTTATCCAGAAATCAGGGGATAACCGCAGGAAGAACAT 6270
Qy 3066 GTGAGCAAAAGGCGCAGCAAGGCCAGGAACCGTAAAGAGCGCGTTCGTCGCGTTTTT 3125
Db 6271 GTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGAGCGCGTTCGTCGCGTTTTT 6330
Qy 3126 CCATAGGCTCCGCCCGCTGACAGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGCGG 3185
Db 6331 CCATAGGCTCCGCCCGCTGACAGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGCGG 6390
Qy 3186 ARAACGACAGGACTATAAGATACCAAGGTTTCCCGCTGGAAGTCCCTCGTCGCTC 3245
Db 6391 ARAACGACAGGACTATAAGATACCAAGGTTTCCCGCTGGAAGTCCCTCGTCGCTC 6450
Qy 3246 TCCTGTTCCACCTCGCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGT 3305
Db 6451 TCCTGTTCCACCTCGCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGT 6510
Qy 3306 GGGCTTTCTCATAGCTCAGCTGATAGTATCTCAGTTCGGTGTAGGTGGTTCGCTCCAA 3365
Db 6511 GGGCTTTCTCAATGCTCAGCTGATAGTATCTCAGTTCGGTGTAGGTGGTTCGCTCCAA 6570
Qy 3366 GCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGGTGCGCTTATCCGCTAACTA 3425
Db 6571 GCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGGTGCGCTTATCCGCTAACTA 6630
Qy 3426 TCGTCTTGAAGTCCAAACCGGTAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAA 3485
Db 6631 TCGTCTTGAAGTCCAAACCGGTAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAA 6690
Qy 3486 CAGGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCCTCAAGTGGTGGCTTAA 3545
Db 6691 CAGGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCCTCAAGTGGTGGCTTAA 6750
Qy 3546 CTACGGCTACACTAGAAAGACAGTATTTGTTGTATCTGCGCTCTGCTGAAGCGAGTTACCTT 3605
Db 6751 CTACGGCTACACTAGAAAGACAGTATTTGTTGTATCTGCGCTCTGCTGAAGCGAGTTACCTT 6810
Qy 3606 CGGAAAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCAACCGCTGGTAGCGGTGGTTT 3665
Db 6811 CGGAAAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCAACCGCTGGTAGCGGTGGTTT 6870
Qy 3666 TTTTGTGTCAGCAGCAGATTACGCGCAGAAAAAGATCTCAAGAGATCCTTTGAT 3725
Db 6871 TTTTGTGTCAGCAGCAGATTACGCGCAGAAAAAGATCTCAAGAGATCCTTTGAT 6930
Qy 3726 CTTTTCTACGGGTCGACGCTCAGTGGACGAAACTCAGTTAAGGATTTTGGTCAT 3785
Db 6931 CTTTTCTACGGGTCGACGCTCAGTGGACGAAACTCAGTTAAGGATTTTGGTCAT 6990
Qy 3786 GAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCTCCCTCCTGCAAGTTTCGGGGGCATG 3845
Db 6991 GAGATTATCAAAAAGGATCTTCACTAGATCCTT----- 7024
Qy 3846 GATGGCGGATAGCGGTCGTGGTTTCTGATGCCACGAGTTTGCACTGCCGCTAGAA 3905
Db 7025 ----- 7024
Qy 3906 CTCGCGAGGTCTGTCAGCCTCAGGCAGCAGCTGAAACCACTCGCGAGGGGATCGAGCCC 3965
Db 7025 -----TT 7026
Qy 3966 GGGGTGGCGAAGAACTCCAGATGAGATCCCGGCTGGAGGATCATTCAGCCGGGCTC 4025
Db 7027 GGGGTGGCGAAGAACTCCAGATGAGATCCCGGCTGGAGGATCATTCAGCCGGGCTC 7086
Qy 4026 CCGGAAAAGATTCCGAGGCCAACCTTTCTAGAGCGCGGCTGGAATCGGAATCTCG 4085
Db 7087 CCGGAAAAGATTCCGAGGCCAACCTTTCTAGAGCGCGGCTGGAATCGGAATCTCG 7146
Qy 4086 TGATGGCAGTTGGCGCTCGTGGTCAATTCGAAACCCCAAGATCCGCTCAGAAG 4145
Db 7147 TGATGGCAGTTGGCGCTCGTGGTCAATTCGAAACCCCAAGATCCGCTCAGAAG 7206

Qy 4146 AACTCGTCAAGAGCGGATAGAAAGCGGATCGCTGCGAATCGGAGCGGATACCGTAA 4205
Db 7207 AACTCGTCAAGAGCGGATAGAAAGCGGATCGCTGCGAATCGGAGCGGATACCGTAA 7266
Qy 4206 AGCAGAGGAGCGGTCAGCCCATTCGCGCGCAAGCTCTTCAGCAATATCACGGTAGCC 4265
Db 7267 AGCAGAGGAGCGGTCAGCCCATTCGCGCGCAAGCTCTTCAGCAATATCACGGTAGCC 7326
Qy 4266 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGCGGCACAGTCGATGAATCCAGAA 4325
Db 7327 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGCGGCACAGTCGATGAATCCAGAA 7386
Qy 4326 AAGCGGCGATTTTCCACATGATATTCGCAAGCAGGATTCGCAATGGGTTCAGCAGAGA 4385
Db 7387 AAGCGGCGATTTTCCACATGATATTCGCAAGCAGGATTCGCAATGGGTTCAGCAGAGA 7446
Qy 4386 TCCCTCGCGCTCGGCGATCGCGCTTGAAGCTCGGCAACAGTTTCGGTGGCGGAGCCCC 4445
Db 7447 TCCCTCGCGCTCGGCGATCGCGCTTGAAGCTCGGCAACAGTTTCGGTGGCGGAGCCCC 7506
Qy 4446 TGATGCTCTTCTGTCAGATCATCTGATCGCAAGACCGGCTTCCATCCGAGTACGTGCT 4505
Db 7507 TGATGCTCTTCTGTCAGATCATCTGATCGCAAGACCGGCTTCCATCCGAGTACGTGCT 7566
Qy 4506 CGCTCGATGCGATGTTTCGTTGGTGTGCAATGGGCGAGGTAGCCGATCAAGCGTATGC 4565
Db 7567 CGCTCGATGCGATGTTTCGTTGGTGTGCAATGGGCGAGGTAGCCGATCAAGCGTATGC 7626
Qy 4566 AGCGCGCGATGTCAGGCAATGATGGATATCTTTCGCGGAGGAGCAAGGTAGATGAC 4625
Db 7627 AGCGCGCGATGTCAGGCAATGATGGATATCTTTCGCGGAGGAGCAAGGTAGATGAC 7686
Qy 4626 AGGAGATCTCGCCCGGCACTTCGCCCAATAGCAGCGAGTCCCTTCCCGCTTCACTGACA 4685
Db 7687 AGGAGATCTCGCCCGGCACTTCGCCCAATAGCAGCGAGTCCCTTCCCGCTTCACTGACA 7746
Qy 4686 ACCTCGAGCAGCTGCGCAAGGAAACCGCTCGTGGCAGCCAGATAGCGCGCTGCGC 4745
Db 7747 ACCTCGAGCAGCTGCGCAAGGAAACCGCTCGTGGCAGCCAGATAGCGCGCTGCGC 7806
Qy 4746 TCGTCTCTGAGTTTCAATTCAGGGCAGCGGACAGGTTCGTTGACAAAAGAACCGGGCGC 4805
Db 7807 TCGTCTCTGAGTTTCAATTCAGGGCAGCGGACAGGTTCGTTGACAAAAGAACCGGGCGC 7866
Qy 4806 CCCTCGCTGACAGCGGCAACCGCGGCAATCAGAGCAGCGGATGTTGTTGTCGCGAG 4865
Db 7867 CCCTCGCTGACAGCGGCAACCGCGGCAATCAGAGCAGCGGATGTTGTTGTCGCGAG 7926
Qy 4866 TCATAGCGAATAGCTCTCCACCGAGCGCGGAGAACCTGGTGGCAATCCATCTGTT 4925
Db 7927 TCATAGCGAATAGCTCTCCACCGAGCGCGGAGAACCTGGTGGCAATCCATCTGTT 7986
Qy 4926 TCAATCATGCGAAACGATCTCATCTGCTCTTTCGATCAGATCTTGTATCCCTCGGCAAT 4985
Db 7987 TCAATCATGCGAAACGATCTCATCTGCTCTTTCGATCAGATCTTGTATCCCTCGGCAAT 8046
Qy 4986 CAGATCTTGGCGGCAAGAAACCATCCAGTTTACCTTTCAGAGGCTTCCGACCTTACCA 5045
Db 8047 CAGATCTTGGCGGCAAGAAACCATCCAGTTTACCTTTCAGAGGCTTCCGACCTTACCA 8106
Qy 5046 GAGGCGCGCCAGCTGGCAATTCGCGTTTCGCTGCTGCTCCATAAAACCGCCAGTCTAGC 5105
Db 8107 GAGGCGCGCCAGCTGGCAATTCGCGTTTCGCTGCTGCTCCATAAAACCGCCAGTCTAGC 8166
Qy 5106 TATCGCCATGTAAGCCCATCTGCAAGTACCTGCTTCTCTTTCGCTTGGCTTGGCTTCCCTT 5165
Db 8167 TATCGCCATGTAAGCCCATCTGCAAGTACCTGCTTCTCTTTCGCTTGGCTTGGCTTCCCTT 8226
Qy 5166 GTCCAGATAGCCAGTGTGATCATTCATCCGGGTGAGCAGCCGTTTCTGGGAGTGGCT 5225
Db 8227 GTCCAGATAGCCAGTGTGATCATTCATCCGGGTGAGCAGCCGTTTCTGGGAGTGGCT 8286

4086 -----TGATGGCAGGTTGGCGTTCGCTTGGTTCGATCATT 4119
Db CAGGTGGGGTCTTTCAATCCCGCTTTTCTGAGGTTGGCGTTCGCTTGGTTCGATCATT 4512
4120 TCGAAGCCCGAGTTCGGCTCAGAAAGTCTGTCAGAAAGGCGATAGAGGCGATCGCT 4179
Db TCGAAGCCCGAGTTCGGCTCAGAAAGTCTGTCAGAAAGGCGATAGAGGCGATCGCT 4452
4180 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCGCCCAA 4239
Db GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCGCCCAA 4392
4240 GCTCTTCAGCAATATACCGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGACACCA 4299
Db GCTCTTCAGCAATATACCGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGACACCA 4332
4300 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAGC 4359
Db GCGCGCCACAGTCGATGAATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAGC 4272
4360 AGGATCGCCATCGGTCACGAGATTCCTCGCGGTTCGGGCATGCGCGCTTGAGCTGG 4419
Db AGGATCGCCATCGGTCACGAGATTCCTCGCGGTTCGGGCATGCGCGCTTGAGCTGG 4212
4420 CGAAGCTTCGGTTCGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 4479
Db CGAAGCTTCGGTTCGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 4152
4480 GACCGGCTTCATCCGAGTACGTCGCTCGATCGATGTTTCGTTGGTGGTGAATG 4539
Db GACCGGCTTCATCCGAGTACGTCGCTCGATCGATGTTTCGTTGGTGGTGAATG 4092
4540 GCGAGTTCGGTTCAGCGATGACGCGCGCATTCGATCGATGATGATGATGATGATGAT 4599
Db GCGAGTTCGGTTCAGCGATGACGCGCGCATTCGATCGATGATGATGATGATGATGAT 4032
4600 TCTCGGAGGAGCAAGTGAGATGACAGGAGATCTCGCCCGCACTTCGCCCAATAGCA 4659
Db TCTCGGAGGAGCAAGTGAGATGACAGGAGATCTCGCCCGCACTTCGCCCAATAGCA 3972
4660 GCGAGTTCCTTCGGTTCAGTACAGCTGCGAGCAGCTGCGGAGGAGCGCCGCTCG 4719
Db GCGAGTTCCTTCGGTTCAGTACAGCTGCGAGCAGCTGCGGAGGAGCGCCGCTCG 3912
4720 TGGCCAGCCAGATAGCGCGCTGCTTCGTCAGTTCATTCAGGCGCACCGGACAGGT 4779
Db TGGCCAGCCAGATAGCGCGCTGCTTCGTCAGTTCATTCAGGCGCACCGGACAGGT 3852
4780 CGGTCTTGAACAAAGAACCGGCGCCCTGCGTGCAGCCGGAACACGCGGATCAG 4839
Db CGGTCTTGAACAAAGAACCGGCGCCCTGCGTGCAGCCGGAACACGCGGATCAG 3792
4840 AGCAGCGATTCGTCGTTGTGCCAGTCATAGCGGATAGCCCTCCACCCAGCGCGG 4899
Db AGCAGCGATTCGTCGTTGTGCCAGTCATAGCGGATAGCCCTCCACCCAGCGCGG 3732
4900 GAGAACCTGGTGCAATCCATCTTTGTTCAATGCGAAACGATCTCATCTCTCTCTT 4959
Db GAGAACCTGGTGCAATCCATCTTTGTTCAATGCGAAACGATCTCATCTCTCTCTT 3672
4960 GATCAGATCTTGATCCCTCGGCATCAGATCTTCGCGGCAAGAGCCATCCAGTTTA 5019
Db GATCAGATCTTGATCCCTCGGCATCAGATCTTCGCGGCAAGAGCCATCCAGTTTA 3612
5020 CTTTTCAGGGCTTCCCAACCTTACAGAGGCGCGCCAGCTGGCAATTCGCGTTCGCTTG 5079
Db CTTTTCAGGGCTTCCCAACCTTACAGAGGCGCGCCAGCTGGCAATTCGCGTTCGCTTG 3552
5080 CTGTCCATAAAGCCCGCTAGCTATGCGCATGTAAGCCCACTGCAAGTACCTGCT 5139
Db CTGTCCATAAAGCCCGCTAGCTATGCGCATGTAAGCCCACTGCAAGTACCTGCT 3492
5140 TTCTCTTTCGCTTGGCTTTTCCCTTGTTCAGATAGCCAGTAGCTGACATTCATCCGGG 5199

Db 3491 TTCTCTTTCGCTTGGCTTTTCCCTTGTTCAGATAGCCAGTAGCTGACATTCATCCGG 3432
Qy 5200 GTACAGCACGCTTTCTGGGACTGGCTTTCTACGTTGTTCCGCTTCTTTAGCAGCCCTTGC 5259
Db 3431 GTACAGCACGCTTTCTGGGACTGGCTTTCTACGTTGTTCCGCTTCTTTAGCAGCCCTTGC 3372
Qy 5260 GCGCTGAGTGTTCGGCAGCGTG 5283
Db 3371 GCGCTGAGTGTTCGGCAGCGTG 3348

RESULT 4
US-09-380-190A-30/c
; Sequence 30, Application US/09380190A
; Patent No. 6410220
; GENERAL INFORMATION:
; APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
; TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES
; THEREOF
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEHLING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09380.190A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/03918
; FILING DATE: 28-FEB-98
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEHLING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 228.00010201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-380-190A-30

Query Match 30.6%; Score 1618.2; DB 4; Length 6561;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 1911; Conservative 0; Mismatches 23; Indels 269; Gaps 1;

Qy 3081 GCMAAGCCAGGAAACCGTAAAGCGCGCTTCGCGGTTTTCATAGCTCCGCCC 3140
Db 2305 GCGAGTGACTCGAGGCGCGCATCTCGCGTTTCTGCGGTTTTCATAGCTCCGCCC 2246
Qy 3141 CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACAGACT 3200
Db 2245 CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACAGACT 2186
Qy 3201 ATAAGATACAGGCGTTTCCCGCTTGGAGCTCCCTCGTGGCTCTCTGTTCCGACCT 3260
Db 2185 ATAAGATACAGGCGTTTCCCGCTTGGAGCTCCCTCGTGGCTCTCTGTTCCGACCT 2126


```

1 / TITLE OF INVENTION: Method of Use
2 /
3 / NUMBER OF SEQUENCES: 8
4 / CORRESPONDENCE ADDRESS:
5 / ADDRESSEE: Lyon & Lyon
6 / STREET: 633 West Fifth Street
7 / STREET: Suite 4700
8 / CITY: Los Angeles
9 / STATE: California
10 / COUNTRY: U.S.A.
11 / ZIP: 90071-2066
12 /
13 / COMPUTER READABLE FORM:
14 / MEDIUM TYPE: 3.5" Diskette, 1.4.4
15 / MEDIUM TYPE: storage
16 / COMPUTER: IBM Compatible
17 / OPERATING SYSTEM: IBM P.C. DOS
18 / SOFTWARE: Word Perfect 5.1
19 / CURRENT APPLICATION DATA:
20 / APPLICATION NUMBER: US/08/472,8
21 / FILING DATE: June 7, 1995
22 / CLASSIFICATION: 435
23 / PRIOR APPLICATION DATA:
24 / APPLICATION NUMBER: 08/209,846
25 / FILING DATE: March 9, 1994
26 / APPLICATION NUMBER: 07/789,919
27 / FILING DATE: No. 5925564member 6
28 / ATTORNEY/AGENT INFORMATION:
29 / NAME: Warburg, Richard J.
30 / REGISTRATION NUMBER: 32,327
31 / REFERENCE/DOCKET NUMBER: 214/21
32 / TELECOMMUNICATION INFORMATION:
33 / TELEPHONE: (213) 489-1600
34 / TELEFAX: (213) 955-0440
35 / TELEX: 67-3510
36 / INFORMATION FOR SEQ ID NO: 8:
37 / SEQUENCE CHARACTERISTICS:
38 / LENGTH: 5707 bases
39 / TYPE: nucleic acid
40 / STRANDEDNESS: double
41 / TOPOLOGY: linear
42 / MOLECULE TYPE: cdna
43 / US-08-472-809B-8

```

	Query Match	30.3%	Score 1602.2;	DB 2;	Length 5707;
	Best Local Similarity	82.8%;	Pred. No. 0;		
	Matches 2021;	Conservative	0;	Mismatches 28;	Indels 392; Gaps 1;
QY	2706	CTAGACGTAATCATGCTCATAGCTGTTTCTCTGTGTGAAATGTTATCCGCTCACAATTC	2765		
Db	3514	CTTGGCGTAATCATGGTCTATAGCTGTTTCTCTGTGTGAAATGTTATCCGCTCACAATTC	3573		
QY	2766	ACACAAATATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGCTA	2825		
Db	3574	ACACAAATATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGCTA	3633		
QY	2826	ACTCACATTAATTCGGTTGCGCTCACATGCCCGCTTTCACATCGGGAAACCTGTCGTGCCA	2885		
Db	3634	ACTCACATTAATTCGGTTGCGCTCACATGCCCGCTTTCACATCGGGAAACCTGTCGTGCCA	3693		
QY	2886	GCTGCATTAATGAATTCGGCCAAACCGCGGGGAGAGCGGTTTGCGTATTGGCGGCTCTTC	2945		
Db	3694	GCTGCATTAATGAATTCGGCCAAACCGCGGGGAGAGCGGTTTGCGTATTGGCGGCTCTTC	3753		
QY	2946	CGCTTCCTCGCTCACTGACTCGCTGCGGTGCGTTCGCTCGGTCGGCGAGCGGTATCAGC	3005		
Db	3754	CGCTTCCTCGCTCACTGACTCGCTGCGGTGCGTTCGCTCGGTCGGCGAGCGGTATCAGC	3813		
QY	3006	TCACCTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGACGGAAGAACAAT	3065		
Db	3814	TCACCTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGACGGAAGAACAAT	3873		
QY	3066	GTGAGCAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGCCCGCTTGTCGGCGTTTTT	3125		
Db	3874	GTGAGCAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGCCCGCTTGTCGGCGTTTTT	3933		

3126	Qy	CCATAGGCTCCGCCCCCTCTGACGAGCATCAAAAATCAACGCTCAAGTCAGAGGTGGC	3185
3934	Db	CCATAGGCTCCGCCCCCTCTGACGAGCATCAAAAATCAACGCTCAAGTCAGAGGTGGC	3993
3186	Qy	AAACCCGACAGGACTATAAGATACCAGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTC	3245
3994	Db	AAACCCGACAGGACTATAAGATACCAGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTC	4053
3246	Qy	TCCTGTTCCGACCCCTCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCCT	3305
4054	Db	TCCTGTTCCGACCCCTCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCCT	4113
3306	Qy	GGCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAA	3365
4114	Db	GGCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAA	4173
3366	Qy	GCTGGGCTGTGCACGNAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACATA	3425
4174	Db	GCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACATA	4233
3426	Qy	TCGTCCTTAGCTCAAAACCCGGTAAGACGACCTTATCGCCATCGCAGCACCACCTGGTAA	3485
4234	Db	TCGTCCTTAGCTCAAAACCCGGTAAGACGACCTTATCGCCATCGCAGCACCACCTGGTAA	4293
3486	Qy	CAGGATTTAGCAGAGCAGGTTATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGGCTAA	3545
4294	Db	CAGGATTTAGCAGAGCAGGTTATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGGCTAA	4353
3546	Qy	CTACGGCTACACTAGAAGAACGATATTGGTATCTCGCTCTGCTGAAGCCAGTTACCTT	3605
4354	Db	CTACGGCTACACTAGAAGAACGATATTGGTATCTCGCTCTGCTGAAGCCAGTTACCTT	4413
3606	Qy	CGGAAAGAGTGTGATGCTCTTGATCCGCAACCAACACCGCTGTAGCGGTGGTTT	3665
4414	Db	CGGAAAGAGTGTGATGCTCTTGATCCGCAACCAACACCGCTGTAGCGGTGGTTT	4473
3666	Qy	TTTTGTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGAT	3725
4474	Db	TTTTGTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGAT	4533
3726	Qy	CTTTTCTACGGGTCACGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGGTCAT	3785
4534	Db	CTTTTCTACGGGTCACGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGGTCAT	4551
3786	Qy	GAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCACTCCTGTCAGTTTCGGGGCATG	3845
4552	Db	GAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCACTCCTGTCAGTTTCGGGGCATG	4551
3846	Qy	GATGCGCGGATAGCGGTCTGTTTCTCGATGTCGACGATTGTCACCTGCCGGTAGAA	3905
4552	Db	GATGCGCGGATAGCGGTCTGTTTCTCGATGTCGACGATTGTCACCTGCCGGTAGAA	4551
3906	Qy	CTCCGCGAGGTCGTCCAGCCTCAGGCAGCAGCTGAACAACTCGGAGGGGATCGAGCCC	3965
4552	Db	CTCCGCGAGGTCGTCCAGCCTCAGGCAGCAGCTGAACAACTCGGAGGGGATCGAGCCC	4551
3966	Qy	GGGTTGGGGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCCGGCGTC	4025
4552	Db	GGGTTGGGGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCCGGCGTC	4551
4026	Qy	CCGAAAAACGATTCGGAAGCCAACTTTTATAGAAGGCGCGGTGGAATCGAAATCTCG	4085
4552	Db	CCGAAAAACGATTCGGAAGCCAACTTTTATAGAAGGCGCGGTGGAATCGAAATCTCG	4551
4086	Qy	TGATGGCAGGTTGGGCGTCGCTTGGTCGGTCAATTTTCGAACCCAGAGTCCCGCTCAGAAG	4145
4552	Db	TGATGGCAGGTTGGGCGTCGCTTGGTCGGTCAATTTTCGAACCCAGAGTCCCGCTCAGAAG	4551
4146	Qy	AACCTGTCAGAGCGCATAGAAGCGGATGCTGTGGAATCGGAGCGGGCGATACCGTAA	4205
4552	Db	AACCTGTCAGAGCGCATAGAAGCGGATGCTGTGGAATCGGAGCGGGCGATACCGTAA	4621

QY 4206 AGCAGGAGGAGGCTCAGCCATTCGCGCCCAAGCTCTTCAGCAATATACCGGTAGCC 4265
DB |||||
QY 4622 AGCAGGAGGAGGCTCAGCCATTCGCGCCCAAGCTCTTCAGCAATATACCGGTAGCC 4681
DB |||||
QY 4266 AACGCTATGTCCTGATAGCGGTCCGCAACACCCAGCGGCCACAGTTCGATGATCCAGAA 4325
DB |||||
QY 4682 AACGCTATGTCCTGATAGCGGTCCGCAACACCCAGCGGCCACAGTTCGATGATCCAGAA 4741
DB |||||
QY 4326 AAGCGGCCATTTTCCACCATGATATTCGCGAGCAGGAGCATGCCATGGTCCAGCAGGA 4385
DB |||||
QY 4742 AAGCGGCCATTTTCCACCATGATATTCGCGAGCAGGAGCATGCCATGGTCCAGCAGGA 4801
DB |||||
QY 4386 TCCTCGCGTGGGAGATGCGCGCTTGAGCTTGGCGAAACAGTTCGGTGGCGAGGCC 4445
DB |||||
QY 4802 TCCTCGCGTGGGAGATGCGCGCTTGAGCTTGGCGAAACAGTTCGGTGGCGAGGCC 4861
DB |||||
QY 4446 TGATGCTCTTCTGTCAGATGATCCTGATCGACAGCAGCGCTTCATCCGAGTACGTC 4505
DB |||||
QY 4862 TGATGCTCTTCTGTCAGATGATCCTGATCGACAGCAGCGCTTCATCCGAGTACGTC 4921
DB |||||
QY 4506 CGCTCGATGCGATGTTTCGCTTGGTGGTGGATGGCAGGAGTAGCGGATCAAGCGTATGC 4565
DB |||||
QY 4922 CGCTCGATGCGATGTTTCGCTTGGTGGTGGATGGCAGGAGTAGCGGATCAAGCGTATGC 4981
DB |||||
QY 4566 AGCGCGGATGTCATCAGCCATGATGATCTTCTCGCAGGAGCAAGGTGATGATGAC 4625
DB |||||
QY 4982 AGCGCGGATGTCATCAGCCATGATGATCTTCTCGCAGGAGCAAGGTGATGATGAC 5041
DB |||||
QY 4626 AGGAGATCTGCGCGGCACTTCGCGCAATAGCAGCAGTCCCTTCGCTTCAGTGACA 4685
DB |||||
QY 5042 AGGAGATCTGCGCGGCACTTCGCGCAATAGCAGCAGTCCCTTCGCTTCAGTGACA 5101
DB |||||
QY 4686 ACCTCGAGCAGCTGCGCAAGCAAGCGCGTCGCGCAGCAGCAGTAGCGCGCTGCC 4745
DB |||||
QY 5102 ACCTCGAGCAGCTGCGCAAGCAAGCGCGTCGCGCAGCAGCAGTAGCGCGCTGCC 5161
DB |||||
QY 4746 TCGTCTCGAGTTCATTCAGGAGCAGCAGAGCTCGGTCTTGCAAAAGAAACCGGCGC 4805
DB |||||
QY 5162 TCGTCTCGAGTTCATTCAGGAGCAGCAGAGCTCGGTCTTGCAAAAGAAACCGGCGC 5221
DB |||||
QY 4806 CCTCGCGCTGACAGCGGCAACCGCGGATCAGAGCAGCGATGCTGTGTGTCGCCAG 4865
DB |||||
QY 5222 CCTCGCGCTGACAGCGGCAACCGCGGATCAGAGCAGCGATGCTGTGTGTCGCCAG 5281
DB |||||
QY 4866 TCATAGCCGATAGCTCTCCACCAAGCGCGGAGAACTCGGTGCAATCCATCTGT 4925
DB |||||
QY 5282 TCATAGCCGATAGCTCTCCACCAAGCGCGGAGAACTCGGTGCAATCCATCTGT 5341
DB |||||
QY 4926 TCAATCATGCGAAACGATCCTCTGCTCTTCATCAGATCTTGATCCCTCGGCGAT 4985
DB |||||
QY 5342 TCAATCATGCGAAACGATCCTCTGCTCTTCATCAGATCTTGATCCCTCGGCGAT 5401
DB |||||
QY 4986 CAGATCTTGGGGGCAAGAAAGCATTCAGTTTACTTTGCGAGGGTTCCTCAACCTTACCA 5045
DB |||||
QY 5402 CAGATCTTGGGGGCAAGAAAGCATTCAGTTTACTTTGCGAGGGTTCCTCAACCTTACCA 5461
DB |||||
QY 5046 GAGGCGGCGCCAGCTGCGCAATTCGCGTTCGCTTGTGTCATATAAAGCGCGAGCTTACG 5105
DB |||||
QY 5462 GAGGCGGCGCCAGCTGCGCAATTCGCGTTCGCTTGTGTCATATAAAGCGCGAGCTTACG 5521
DB |||||
QY 5106 TATCGCATGTAGCCCACTGCAAGTACCTGCTTCTCTT 5146
DB |||||
QY 5522 AACTGTTGGAGGGGATCGGTGGGGCTCTTCGCTATT 5562
DB |||||

RESULT 6

US-08-948-378A-7/c
; Sequence 7, Application US/08948378A
; Patent No. 6013258
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicx, Roman M.
; APPLICANT: Collins, Edward J.

APPLICANT: Hedley, Mary Lynn
TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM
THE HPV E7 PROTEIN
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Fast-Seq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,378A
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 08191/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: pBIOTOPHPV
US-08-948-378A-7

Query Match 29.7%; Score 1571.6; DB 3; Length 4665;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 114; Indels 180; Gaps 3;
QY 3008 ACTCAAGGCGGTATACGGTTATCCAGATCAGGGGATACGCGAGGAAGAACATGT 3067
DB |||||
QY 2624 AATCATGCGCGTATACGGTTATCCAGATCAGGGGATACGCGAGGAAGAACATGT 2565
DB |||||
QY 3068 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGGTTCCTGGCGTTTTC 3127
DB |||||
QY 2564 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGGTTCCTGGCGTTTTC 2505
DB |||||
QY 3128 ATAGCTCGCGCCCGCTGACGAGCATCAAAAATCGACCTCAAGTCAGAGTGGCGAA 3187
DB |||||
QY 2504 ATAGCTCGCGCCCGCTGACGAGCATCAAAAATCGACCTCAAGTCAGAGTGGCGAA 2445
DB |||||
QY 3188 ACCCGACAGGACTATAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 3247
DB |||||
QY 2444 ACCCGACAGGACTATAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2385
DB |||||
QY 3248 CTGTTCCGACCGCTCGCGTTACCGGATACCTGTCGGCTTCTCCCTTCGGGAAGGTGG 3307
DB |||||
QY 2384 CTGTTCCGACCGCTCGCGTTACCGGATACCTGTCGGCTTCTCCCTTCGGGAAGGTGG 2325
DB |||||
QY 3308 CGCTTTCTCATAGCTCAGCTGATAGTATCTCAGTTCCGTTAGTGTCTTCGTTCCGAAGC 3367
DB |||||
QY 2324 CGCTTTCTCATAGCTCAGCTGATAGTATCTCAGTTCCGTTAGTGTCTTCGTTCCGAAGC 2265
DB |||||
QY 3368 TGGGCTGTGTGCAAGAACCCCGCTTCAGCGGACCGCTCGGCTTATCCGTAATATC 3427
DB |||||
QY 2264 TGGGCTGTGTGCAAGAACCCCGCTTCAGCGGACCGCTCGGCTTATCCGTAATATC 2205
DB |||||
QY 3428 GTCTTGTAGTCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACA 3487
DB |||||

2204	Db	GTCTTTGAGTCCAAACCCGGTAAAGACACGAGCTTATCCGCTACTCGCAGCAGCCACTGGTAAACA	2145
3488	Qy	GGATTAGCAGACGAGGATGTAGCGCGTGTCTACAGAGTTCTTGAAGTGGTGGCCTAACT	3547
2144	Db	GGATTAGCAGACGAGGATGTAGCGCGTGTCTACAGAGTTCTTGAAGTGGTGGCCTAACT	2085
3548	Qy	ACGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTATACCTTCG	3607
2084	Db	ACGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTATACCTTCG	2025
3608	Qy	GAATAAGAGTTGGTAGCTTTCATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT	3667
2024	Db	GAATAAGAGTTGGTAGCTTTCATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT	1965
3668	Qy	TTGTTTGAAGCAGCAGATTAACGGCAGAAAAAAGATCTCAAGAAAGATCCTTTGATCT	3727
1964	Db	TTGTTTGAAGCAGCAGATTAACGGCAGAAAAAAGATCTCAAGAAAGATCCTTTGATCT	1905
3728	Qy	TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTTAAGGGATTTTGGTCATGA	3787
1904	Db	TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTTAAGGGATTTTGGTCATGA	1845
3788	Qy	GAATATCGTCGACCAAGCGGCCATCGTGCCT-----	3819
1844	Db	GAATATCAAAAAGGATCTTTCAGCTAGTCTTTTAAATTTAAAAATGAAGTTTTTAAATCAA	1785
3820	Qy	-----CCCCACTCTGTCAGTTTCGGGG	3840
1784	Db	TCATAAGTATATATAGTAACTGAGGCTATGGCAGGGCTCGCGCCGACGTTGGCTG	1725
3841	Qy	GCATGATCGCGGATAGCGCTGTCTGGTTTCTGTGATCCGACCGGATTTGACCTGCCGG	3900
1724	Db	CGAGCCCTGGGCTTTCACCGGAATTTGGGGGTGGGGTGGGAAAGGAAGAAACGCGGG	1665
3901	Qy	-----TAGAACTCGCGAGTCTGTCAGCCTCAGGCAGCAGCTGAACCAAC	3946
1664	Db	CGTATTTGGCCCCAATGGGTTCTCGTGGGGTATCGACAGTGCACGCCCTGGACCGNA	1605
3947	Qy	TCGCGAGGGGATCGA-----	3961
1604	Db	CCCCGGTTTATGAACAAACGCCAACCGTGGCTTTATCTGTCTTTTATTGGCG	1545
3962	Qy	-----GCCCGG	3967
1544	Db	TCATAGCGCGGTTTCCTTCGSPATTGTCTCTTCGTTGTTTCAGTTAGCTCCCCCTAG	1485
3968	Qy	GGTGGGGAAGAACTCCAGCATCAGATCCCGCGCTGGAGGATCATCCAGCGCGGCTCC	4027
1484	Db	GGTGGGGAAGAACTCCAGCATCAGATCCCGCGCTGGAGGATCATCCAGCGCGGCTCC	1425
4028	Qy	GGAAAAAGATTCGGAAGCCAACTTTTCATAGAAGCGCGGCTGGGAATCGAATCTCGTG	4087
1424	Db	GGAAAAAGATTCGGAAGCCAACTTTTCATAGAAGCGCGGCTGGGAATCGAATCTCGTG	1365
4088	Qy	ATGCGAGTTGGGGCTGCTTGTTCGTCTATTTTGAACCCAGAGTCCCGCTCAGAAGAA	4147
1364	Db	ATGCGAGTTGGGGCTGCTTGTTCGTCTATTTTGAACCCAGAGTCCCGCTCAGAAGAA	1305
4148	Qy	CTCGTCAAGAAGCGCATAGAAGCGATGCGCTCGAATTCGGAGCGCGCATACCGTAAAG	4207
1304	Db	CTCGTCAAGAAGCGCATAGAAGCGATGCGCTCGAATTCGGAGCGCGCATACCGTAAAG	1245
4208	Qy	CACAGGAAGCGGTACGCCCATTTCCGCGCCAGAGCTCTTCAGCAATATCACGGGTAGCCAA	4267
1244	Db	CACAGGAAGCGGTACGCCCATTTCCGCGCCAGAGCTCTTCAGCAATATCACGGGTAGCCAA	1185
4268	Qy	CGCTATGTCCTGATAGGGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAA	4327
1184	Db	CGCTATGTCCTGATAGGGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAA	1125
4328	Qy	CGGGCCATTTTCCACATGATATTCGGCAGCAGGCATTCGCCATGGGTTCACGAGATC	4387
1124	Db	CGGGCCATTTTCCACATGATATTCGGCAGCAGGCATTCGCCATGGGTTCACGAGATC	1065

Qy	4388	CTCGCGTGGGCAATCGCGCCTTGGAGCCTGGCGGAACAGTTTCGGCTTGGCGCAGCCCTG	4444
Db	1064	CTCGCGTGGGCAATCGCTTGAAGCTTGGCGAACAGTTTCGGCTGGCGGAGCCCTG	1005
Qy	4448	ATGCTCTTCGTCCAGATCATCTTGATCGAACAGACCGGCTTCATCCGAGTAGTGTCTCG	4507
Db	1004	ATGCTCTTCGTCCAGATCATCTTGATCGAACAGACCGGCTTCATCCGAGTAGTGTCTCG	945
Qy	4508	CTCGATCGGATGTTTCGCTTGGTGGTTCGAATGGCAGAGTAGCCGGATCAACGGTATGCAG	4567
Db	944	CTCGATCGGATGTTTCGCTTGGTGGTTCGAATGGCAGAGTAGCCGGATCAACGGTATGCAG	888
Qy	4568	CCGCGCATTTGCATCAGCCATGATGGAATCTTCTCGGCAAGGACAAAGTGGATGACAG	4627
Db	884	CCGCGCATTTGCATCAGCCATGATGGAATCTTCTCGGCAAGGACAAAGTGGATGACAG	825
Qy	4628	GAGATCTGTCGCCCGGCATCTTGGCCCAATAGCAGCCAGTCCCTCTCCCGCTTCAGTGACAAC	4687
Db	824	GAGATCTGTCGCCCGGCATCTTGGCCCAATAGCAGCCAGTCCCTCTCCCGCTTCAGTGACAAC	765
Qy	4688	GTCCAGACACAGCTGCGCAAGAAACGCCCGCTGTGGCCAGCCACGATAGCCGCTGCCTC	4747
Db	764	GTCCAGACACAGCTGCGCAAGAAACGCCCGCTGTGGCCAGCCACGATAGCCGCTGCCTC	705
Qy	4748	GTCCTGCAGTTTCATTTCAGGCGCACCGGACAGTTCGGTCTTGACAAAAGAAACCGCGCGCCC	4807
Db	704	GTCCTGCAGTTTCATTTCAGGCGCACCGGACAGTTCGGTCTTGACAAAAGAAACCGCGCGCCC	645
Qy	4808	CTCGCTGCACAGCCGGAAACACGGCGGCATCAGAGCAGCCGATTCCTCTGTTGTGCCAGTC	4867
Db	644	CTCGCTGCACAGCCGGAAACACGGCGGCATCAGAGCAGCCGATTCCTCTGTTGTGCCAGTC	585
Qy	4868	ATAGCCGAATAGCCTCTCCACCCAAAGCGCGCGGAGAACCTTCGCTGCAATCCATCTTGTTC	4927
Db	584	ATAGCCGAATAGCCTCTCCACCCAAAGCGCGGAGAACCTTCGCTGCAATCCATCTTGTTC	525
Qy	4928	AATCATGCGAAACAGATCTCTCATCTGTCTCTTGATCAGATCTTG	4971
Db	524	AATCATGCGAAACAGATCTCTCATCTGTCTCTTGATCAGATCTTG	481

RESULT 7

US-09-169-425C-7/c
; Sequence 7, Application US/09169425C
; Patent No. 6183746
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chiczo, Roman M.
; APPLICANT: Collins, Edward J.
; APPLICANT: Hedley, Mary Lynn
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM THE HPV E7
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/169,425C
; FILING DATE: 09-OCT-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/061,657
; FILING DATE: 09-OCT-1997
; ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34, 819
REFERENCE/DOCKET NUMBER: 08191/004002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-543-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4665 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-169-425C-7

Query Match 29.7%; Score 1571.6; DB 3; Length 4665;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 114; Indels 180; Gaps 3;

QY	3008	ACTCAAGCGGTAATACGGTTATCCACAGATCAGGGATPACGCGAGGAAGAACATGT	3067
DB	2624	AATGATCGGTAATACGGTTATCCACAGATCAGGGATPACGCGAGGAAGAACATGT	2565
QY	3058	GAGCAAAAGCCAGCAAAAGCCAGGACCGTAAGAGCCGCGTTCGTGCGTTTTTC	3127
DB	2564	GAGCAAAAGCCAGCAAAAGCCAGGACCGTAAGAGCCGCGTTCGTGCGTTTTTC	2505
QY	3128	ATAGGCTCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAA	3187
DB	2504	ATAGGCTCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAA	2445
QY	3188	ACCGACAGGACTAAGATACAGGCGTTTCCCTTGGAGCTCCCTCGTGGCTCTC	3247
DB	2444	ACCGACAGGACTAAGATACAGGCGTTTCCCTTGGAGCTCCCTCGTGGCTCTC	2385
QY	3248	CTGTTCCGACCTCGCGTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG	3307
DB	2384	CTGTTCCGACCTCGCGTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG	2325
QY	3308	CGCTTCTCATAGCTACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCCAAGC	3367
DB	2324	CGCTTCTCATAGCTACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCCAAGC	2255
QY	3368	TGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTTATCCGGTAAGTATC	3427
DB	2264	TGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTTATCCGGTAAGTATC	2205
QY	3428	GTCTTGAGTCCAAACCGGTAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAACA	3487
DB	2204	GTCTTGAGTCCAAACCGGTAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAACA	2145
QY	3488	GGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT	3547
DB	2144	GGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT	2085
QY	3548	ACGGCTACACTAGAAAGACAGTATTTGGTATCTGGCTCTGTGAAAGCAAGTACCTTCG	3607
DB	2084	ACGGCTACACTAGAAAGACAGTATTTGGTATCTGGCTCTGTGAAAGCAAGTACCTTCG	2025
QY	3608	GAAAAGAGTTGTAGCTCTTGTATCGGCAAAACCAACACCGCTGTAGCGGTGTTTTT	3667
DB	2024	GAAAAGAGTTGTAGCTCTTGTATCGGCAAAACCAACCGCTGTAGCGGTGTTTTT	1965
QY	3668	TTGTTTGAAGCAGCAGATTTACCGCGAGAAAAGGATCTCAAGAGATCTTTGATCT	3727
DB	1964	TTGTTTGAAGCAGCAGATTTACCGCGAGAAAAGGATCTCAAGAGATCTTTGATCT	1905
QY	3728	TTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGATTTGGTCAAGA	3787
DB	1904	TTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGATTTGGTCAAGA	1845
QY	3788	GATTATCGTCCAAAGCGGCATCGTGCT-----	3819

DB	1844	GATTATCAAAAGGATCTTACCTAGATCCTTTAAATATAAAATGAAGTTTTAAATCAA	1785
QY	3820	-----CCCACTCTCTCAGTTTCGGG	3840
DB	1784	TCTAAGATATATAGAGTAACCTGAGGCTATGGCAGGCGCTCGCGCCGACGTTGGCTG	1725
QY	3841	GCATGATGCGCGATAGCCGCTGCTGTTTCCTGATCCGACGGATTTGCACCTGCGG	3900
DB	1724	CGAGCCCTGGGCTTACCCGAACTTGGGGGTGGGGTGGGAAAGAAAGAAACCGCGG	1665
QY	3901	-----TAGAACTCCGCGAGTCTGTCAGCCTCAGGCGAGCAGCTGAACCAAC	3946
DB	1664	CGTATTGGCCCAATGGGCTCTCGTGGGTATCGACAGTGCAGCCCTGGGACGAA	1605
QY	3947	TCGCGAGGGGATCGA-----	3961
DB	1604	CCCCCGGTTTATGAACAAACGACCAACCGCTGCGTTTTATTCTGTCTTTTATTGCGG	1545
QY	3962	-----GCCCCG	3967
DB	1544	TCATAGCGGGTTCCTTCCGCTATGCTCTCTCCGTGTTTCAGTTAGCTCCCTCCCTAG	1485
QY	3968	GGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGGAGGATCATCCAGCGCGCTGCC	4027
DB	1484	GGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGGAGGATCATCCAGCGCGCTGCC	1425
QY	4028	GGAAACGATTCGGAAGCCCAACCTTCATAGAAGCGCGCTGGAATCGAAATCTCGTG	4087
DB	1424	GGAAACGATTCGGAAGCCCAACCTTCATAGAAGCGCGCTGGAATCGAAATCTCGTG	1365
QY	4088	ATGGCAGGTTGGCGCTCGCTTCGTCATTTCGAACCCCAAGAGTCCGCTCAGAGAA	4147
DB	1364	ATGGCAGGTTGGCGCTCGCTTCGTCATTTCGAACCCCAAGAGTCCGCTCAGAGAA	1305
QY	4148	CTCGTCAAGAGCGGATAGAGCGATGCGCTCGGAATCGGAGCGGCGATACCGTAAAG	4207
DB	1304	CTCGTCAAGAGCGGATAGAGCGGATGCGCTCGGAATCGGAGCGGCGATACCGTAAAG	1245
QY	4208	CACGAGAAAGCGGTGAGCCCATTCGCGCGCAAGAGCTCTTCAGCAATATCAAGGATGACCAA	4267
DB	1244	CACGAGAAAGCGGTGAGCCCATTCGCGCGCAAGAGCTCTTCAGCAATATCAAGGATGACCAA	1185
QY	4268	CGCTATGCTGTAGTAGCGGTCCGCAACCGAGCGGCGGCGATGATGAATCCGAA	4327
DB	1184	CGCTATGCTGTAGTAGCGGTCCGCAACCGAGCGGCGGCGATGATGAATCCGAA	1125
QY	4328	GCGCCCATTTTCCACCATGATTTCCGCAAGCAGGCGATCCCATGGGTACGACGAGATC	4387
DB	1124	GCGCCCATTTTCCACCATGATTTCCGCAAGCAGGCGATCCCATGGGTACGACGAGATC	1065
QY	4388	CTCGCGGTGGGCGATGCGCGCTTGAAGCTTGGCGAAGCAGTTCGGCTGGCGCGGCGCTG	4447
DB	1064	CTCGCGGTGGGCGATGCGCGCTTGAAGCTTGGCGAAGCAGTTCGGCTGGCGCGGCGCTG	1005
QY	4448	ATGCTCTTCGTCCAGATCATCTGATCGCAACGCGCTTCCATCCGAGTACGTCCTCG	4507
DB	1004	ATGCTCTTCGTCCAGATCATCTGATCGCAACGCGCTTCCATCCGAGTACGTCCTCG	945
QY	4508	CTCGATCGGATGTTTCGTTGGTTCGAATGGGCGAGGTAGCCGATCAAGCGTATGCG	4567
DB	944	CTCGATCGGATGTTTCGTTGGTTCGAATGGGCGAGGTAGCCGATCAAGCGTATGCG	885
QY	4568	CCGCGCATTTGCATCAGCCATGATGATCTTCTCGGAGGAGCAAGGTGAGTACACAG	4627
DB	884	CCGCGCATTTGCATCAGCCATGATGATCTTCTCGGAGGAGCAAGGTGAGTACACAG	825
QY	4628	GAGATCTGCCCCGCACTTCGCCCAATAGCAGCTCCCTTCCTCCGCTTCAGTGACAAC	4687
DB	824	GAGATCTGCCCCGCACTTCGCCCAATAGCAGCTCCCTTCCTCCGCTTCAGTGACAAC	765
QY	4688	GTCAGACAGCTGCGCAAGAACGCGCTGCGGCGAGCAGGATAGCGCGCTGCTC	4747

Db 764 GTCGAGCAGACGTGGCAAGAACGCCCGCTGCTGGCCAGCCACGATAGCGCGCTGCCTC 705
Qy 4748 GTCCTGCAGTTTCATTCAGGCGACCGGACAGGTGCTTTCACAAAAGAACCGGCGGCC 4807
Db 704 GTCCTGCAGTTTCATTCAGGCGACCGGACAGGTGCTTTCACAAAAGAACCGGCGGCC 645
Qy 4808 CTGGGCTGACAGCGCGAACACCGGCGCATCAGAGCAGCGGATGCTGTTGTGCCCAGTC 4867
Db 644 CTGGGCTGACAGCGCGAACACCGGCGCATCAGAGCAGCGGATGCTGTTGTGCCCAGTC 585
Qy 4868 ATAGCGGATAGCTCTCCACCAAGCGGCGGAGAACCTGCGTGAATCCTTTGTTTC 4927
Db 584 ATAGCGGATAGCTCTCCACCAAGCGGCGGAGAACCTGCGTGAATCCTTTGTTTC 525
Qy 4928 AATCATGGAACGATCCTCATCCTGCTCTCTGATCAGATCTTG 4971
Db 524 AATCATGGAACGATCCTCATCCTGCTCTCTGATCAGATCTTG 481

RESULT 8

US-09-759-960-7/c
; Sequence 7, Application US/09759960
; Patent No. 6582704
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicz, Roman M.
; APPLICANT: Collins, Edward J.
; APPLICANT: Hedley, Mary Lynn
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM THE HPV E7
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09759,960
; APPLICATION NUMBER: US/09759,960

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/169,425
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/004002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-543-8906
; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-759-960-7

Query Match 29.7%; Score 1571.6; DB 4; Length 4665;

Best Local Similarity 86.3%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 114; Indels 180; Gaps 3;

Qy 3008 ACTCAAGCGGTAAATCGGTATCCACAGATCAGGGGATACGAGGAAAGAACATGT 3067
Db 2624 AATGATCGCGGTAAATCGGTATCCACAGATCAGGGGATACGAGGAAAGAACATGT 2565

Qy 3068 GAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 3127
Db 2564 GAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 2505
Qy 3128 ATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 3187
Db 2504 ATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 2445
Qy 3188 ACCGAGCAGACTATAAAGATACGAGCGTTTCCGCCCTTCCGAAAGCTCCCTCGCTGCGTCTC 3247
Db 2444 ACCGAGCAGACTATAAAGATACGAGCGTTTCCGCCCTTCCGAAAGCTCCCTCGCTGCGTCTC 2385
Qy 3248 CTGTTCCGACCTGCCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGG 3307
Db 2384 CTGTTCCGACCTGCCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGG 2325
Qy 3308 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGTGTAGTCTGCTCTCAAGC 3367
Db 2324 CGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGTGTAGTCTGCTCTCAAGC 2265
Qy 3368 TGCGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCCTTATCCGTAACCTATC 3427
Db 2264 TGCGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCCTTATCCGTAACCTATC 2205
Qy 3428 GTCTTGAGTCCAAACCGGTAAAGACGACTTATCGCACTGCGCAGCAGCCACTGGTAACA 3487
Db 2204 GTCTTGAGTCCAAACCGGTAAAGACGACTTATCGCACTGCGCAGCAGCCACTGGTAACA 2145
Qy 3488 GATTAGCAGAGCGAGGTATGTAGCGGTGTGTACAGAGTTCTTGAAGTGGTGGCCTAACT 3547
Db 2144 GATTAGCAGAGCGAGGTATGTAGCGGTGTGTACAGAGTTCTTGAAGTGGTGGCCTAACT 2085
Qy 3548 ACGGCTACACTAGAGAAACAGTATTTGGTATCTGCTGCTGCTGAGCAGCTTACCTTCG 3607
Db 2084 ACGGCTACACTAGAGAAACAGTATTTGGTATCTGCTGCTGCTGAGCAGCTTACCTTCG 2025
Qy 3608 GAAAAAGAGTTGTTAGTCTCTTGATCCGGCAAAACCAACCCGCTGCTGAGCGTGGTTTTT 3667
Db 2024 GAAAAAGAGTTGTTAGTCTCTTGATCCGGCAAAACCAACCCGCTGCTGAGCGTGGTTTTT 1965
Qy 3668 TTGTTTGCAGCAGCAGATTCGCGCAGAAAAGGATCTCAAGAGATCTTTGATCT 3727
Db 1964 TTGTTTGCAGCAGCAGATTCGCGCAGAAAAGGATCTCAAGAGATCTTTGATCT 1905
Qy 3728 TTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATGA 3787
Db 1904 TTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATGA 1845
Qy 3788 GATTATCGTCGACCAAGCGGCATCGTGCCT----- 3819
Db 1844 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAGTAAAGTTTAAATCAA 1785
Qy 3820 -----CCCCACTCTCGAGTTTCGGGG 3840
Db 1784 TCTAAAGTATATATGAGTAACCTGAGGCTATGGCAGGGGCTGCGCCCGCGAGTTGGCTG 1725
Qy 3841 GCATGGATGCGGATAGCGCTGCTGCTTCTGATGCGCAGCGATTTGACCTGCGG 3900
Db 1724 CGAGCCCTGGGCGCTTCCCGGAACTTGGGGGTGGGGGAAAGGAAAGAACCGCGG 1665
Qy 3901 -----TAGAACTCCGAGGTCTCCAGCCTCAGGCGAGCTGAAACCAAC 3946
Db 1664 CGTATTGGCCCCAAATGGGGTCTCGGTGGGTATCGACAGAGTCCAGCCCTGGGACCGAA 1605
Qy 3947 TCCGAGGGGATCGA----- 3961
Db 1604 CCCCCTTTATGAACAAACGACCCCAACCGCTGCTTTTATCTGCTCTTTTATGCGG 1545
Qy 3962 -----GCCCG 3967
Db 1544 TCATAGCGGGTTCCTTCGGGTATTGCTCCTTCGGTGTTCAGTTAGCTTCCCTCCCTAG 1485

QY 3428 GTCTTGGTCCACCGGTAAGACACGACTTATCCGCACTGCGCAGCGCCACTGGTAACA 3487
 Db 2315 GTCTTGGTCCACCGGTAAGACACGACTTATCCGCACTGCGCAGCGCCACTGGTAACA 2374
 QY 3488 GGATTAGCAGCGAGGTATAGCGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACT 3547
 Db 2375 GGATTAGCAGCGAGGTATAGCGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACT 2434
 QY 3548 ACGGCTACACTAGAACAGATATTGGTATCTCGGCTCTGCTGAAGCCAGTACCTTCG 3607
 Db 2435 ACGGCTACACTAGAACAGATATTGGTATCTCGGCTCTGCTGAAGCCAGTACCTTCG 2494
 QY 3608 GAAAAAGAGTGTGACTCTTGAATCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTTT 3667
 Db 2495 GAAAAAGAGTGTGACTCTTGAATCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTTT 2554
 QY 3668 TTGTTTTCGAAAGCAGAGATTAGCGCAGAAAAAAGAGATCTCAAGAGATCTTTGATCT 3727
 Db 2555 TTGTTTTCGAAAGCAGAGATTAGCGCAGAAAAAAGAGATCTCAAGAGATCTTTGATCT 2614
 QY 3728 TTTCTACGGGCTGTGAGCTCAGTGGAGCAAACTCAGTTAGGATTTTGTGTCATGA 3787
 Db 2615 TTTCTACGGGCTGTGAGCTCAGTGGAGCAAACTCAGTTAGGATTTTGTGTCATGA 2674
 QY 3788 GATTATCGTTCGACCAAGCGGCATCGTGCCT- 3819
 Db 2675 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAA 2734
 QY 3820 - 3840
 Db 2735 TCTAAAGTATATAGTAACTGAGGCTATGGCAGGCGCTGCGCCCGGACGTTGGCTG 2794
 QY 3841 GCATGATGCGCGGATAGCGCTGTGTTTCTCGATCGCGACGATTTGCATCTGCGG 3900
 Db 2795 CGAGCCCTGGCTTCAACCGGACTTGGGGGTGGGGTGGGAAAGAAAGAAACGCGGG 2854
 QY 3901 - 3946
 Db 2855 CGTATTGGCCCCAATGGGCTCGTGGGGGTATGACAGAGTGCAGCCCTGGGACCGAA 2914
 QY 3947 TCAGAGGGGATCGA- 3961
 Db 2915 CCGCGGTTTATGAACAAAGACCAACACCGTGGTATTTATCTGCTTTTATTTGCGG 2974
 QY 3962 - 3967
 Db 2975 TCTAGCGCGGTTCTTCGGTATTTGCTCTCGTGGTTCAGTTAGCTTCCCTCCCTAG 3034
 QY 3968 GGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGCTCC 4027
 Db 3035 GGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGCTCC 3094
 QY 4028 GGAAGACGATTCGAAGCCCACTTTCTAGAGCGCGGTGGAATCGAAATCTGCTG 4087
 Db 3095 GGAAGACGATTCGAAGCCCACTTTCTAGAGCGCGGTGGAATCGAAATCTGCTG 3154
 QY 4088 ATGSCAGTTGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4147
 Db 3155 ATGSCAGTTGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3214
 QY 4148 CTGCTCAAGAGCGGATAGAGGAGTGGCTGCGAATCGGAGCGGCGGATACCGTAAAG 4207
 Db 3215 CTGCTCAAGAGCGGATAGAGGAGTGGCTGCGAATCGGAGCGGCGGATACCGTAAAG 3274
 QY 4208 CACGAGGAGCGGTACGCCATTTCCGCGCAAGCTTTTCCAGCAATATCAACGGGTAGCAA 4267
 Db 3275 CACGAGGAGCGGTACGCCATTTCCGCGCAAGCTTTTCCAGCAATATCAACGGGTAGCAA 3334
 QY 4268 CGCTATGCTCTGATAGCGGTCCGCAACCCAGCGCGGCACAGTGCATGATTCAGAGAAA 4327
 Db 3335 CGCTATGCTCTGATAGCGGTCCGCAACCCAGCGCGGCACAGTGCATGATTCAGAGAAA 3394
 QY 4328 CGCGCCATTTTCCACCATGATATTCGGCAAGCAGGCGATCGCCATGGGTCAAGCAGATC 4387

Db 3395 CGGCCATTTTCCACCATGATATTTCGCAAGCAGGATCGCCATGSGTCACGAGATC 3454
 QY 4388 CTCGCGCTGGGATGCGCGCTTGAGCTTGGCGAACAGTTTCGGCTGGCGGAGCCCTG 4447
 Db 3495 CTCGCGCTGGGATGCGCGCTTGAGCTTGGCGAACAGTTTCGGCTGGCGGAGCCCTG 3514
 QY 4448 ATGCTCTTTCGTCAGAGATCATCTGATCGAACAGCGGCTTCCATCCGAGTACGTGCTG 4507
 Db 3515 ATGCTCTTTCGTCAGAGATCATCTGATCGAACAGCGGCTTCCATCCGAGTACGTGCTG 3574
 QY 4508 CTCGATGCGATTTTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4567
 Db 3575 CTCGATGCGATTTTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3634
 QY 4568 CCGCGCATTCATCAGCAGCATGATGGATCTTTCTCGGAGGAGCAAGGTGAGATGACAG 4627
 Db 3635 CCGCGCATTCATCAGCAGCATGATGGATCTTTCTCGGAGGAGCAAGGTGAGATGACAG 3694
 QY 4628 GAGATCTTCCCGCGCATCTTCCGCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAA 4687
 Db 3695 GAGATCTTCCCGCGCATCTTCCGCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAA 3754
 QY 4688 GTCGAGCAGCTGCGCAAGAAACCGCGCTGCTGGCAGCAGATAGCGCGCTGCTC 4747
 Db 3755 GTCGAGCAGCTGCGCAAGAAACCGCGCTGCTGGCAGCAGCAGATAGCGCGCTGCTC 3814
 QY 4748 GTCGAGCAGCTGTCAGGCGCACCGGACAGTCCGCTTTCGACAAAAGAACCGCGGCGCC 4807
 Db 3815 GTCGAGCAGCTTTCAGGCGCACCGGACAGTCCGCTTTCGACAAAAGAACCGCGGCGCC 3874
 QY 4808 CTGCGCTCAGCAGCGGAAACCGCGGCGCATCAGACAGCGGATTTCTGTTGTCGCGAGTC 4867
 Db 3875 CTGCGCTCAGCAGCGGAAACCGCGGCGCATCAGACAGCGGATTTCTGTTGTCGCGAGTC 3934
 QY 4868 ATAGCGGATAGCTTCTCCACCGCGGCGGAGAACCTTGGCTGCAATCCATCTTCTTC 4927
 Db 3935 ATAGCGGATAGCTTCTCCACCGCGGCGGAGAACCTTGGCTGCAATCCATCTTCTTC 3994
 QY 4928 AATCATGGAACGATCT 4971
 Db 3995 AATCATGGAACGATCT 4038

RESULT 10

US-09-533-220A-4/c
 ; Sequence 4, Application US/09533220A
 ; Patent No. 6406908
 ; GENERAL INFORMATION:
 ; APPLICANT: McIntyre, Peter
 ; APPLICANT: James, Iain Fraser
 ; TITLE OF INVENTION: Human Vanilloid Receptor
 ; FILE REFERENCE: 4-30875A
 ; CURRENT APPLICATION NUMBER: US/09/533,220A
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR APPLICATION NUMBER: UNITED KINGDOM 9907097.1
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 1.30
 ; SEQ ID NO 4
 ; LENGTH: 4886
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-533-220A-4

Query Match 29.7%; Score 1570; DB 4; Length 4886;
 Best Local Similarity 86.2%; Pred. No. 0;
 Matches 1849; Conservative 0; Mismatches 115; Indels 180; Gaps 3;
 QY 3008 ACTCAAGCGGTAATACGGTTATCCACAGATCAGGGGATACCGAGGAAGAACATGT 3067
 Db 4295 AATGATGCGGTAATACGGTTATCCACAGATCAGGGGATACCGAGGAAGAACATGT 4236

3068	QY	GAGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGAGGCGCGTGTGCGGTTTTTCC	3127
4235	DB	GAGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGAGGCGCGTGTGCGGTTTTTCC	4176
3128	QY	ATAGGCTCGCCCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAA	3187
4175	DB	ATAGGCTCGCCCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAA	4116
3188	QY	ACCGCAGAGGACTATAAGAGATACCAAGGCGTTTTCCCGCTGGAAAGCTCCCTCGTGGCTCTC	3247
4115	DB	ACCGCAGAGGACTATAAGAGATACCAAGGCGTTTTCCCGCTGGAAAGCTCCCTCGTGGCTCTC	4056
3248	QY	CTGTTCGACCGCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCTTCGGGAGCGTGG	3307
4055	DB	CTGTTCGACCGCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCTTCGGGAGCGTGG	3996
3308	QY	CGTTTTCTCATAGCTCACGCTGTAGTGATCTCAGTTTCGGTGTAGTGTGCTCCAAAGC	3367
3995	DB	CGTTTTCTCATAGCTCACGCTGTAGTGATCTCAGTTTCGGTGTAGTGTGCTCCAAAGC	3936
3368	QY	TGGGCTGTGTCACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCCGTAACTATC	3427
3935	DB	TGGGCTGTGTCACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCCGTAACTATC	3876
3428	QY	GTCTTGAGTCCAAACCGGCTAAGACAGCACTTATCGCCACTGGCAGCAGCACTGGTAACA	3487
3875	DB	GTCTTGAGTCCAAACCGGCTAAGACAGCACTTATCGCCACTGGCAGCAGCACTGGTAACA	3816
3488	QY	GGATTAGCAGGAGGATGTATGAGGCGGTGCTACAGAGTCTTGAGTGGTGCCCTACT	3547
3815	DB	GGATTAGCAGGAGGATGTATGAGGCGGTGCTACAGAGTCTTGAGTGGTGCCCTACT	3756
3548	QY	ACGGCTACACTAGAAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTCG	3607
3755	DB	ACGGCTACACTAGAAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTCG	3696
3608	QY	GAAAAAGAGTTGGTAGTCTTGATTCGGGCAAAACAAACCCGCTGGTAGGGTGGTTTT	3667
3695	DB	GAAAAAGAGTTGGTAGTCTTGATTCGGGCAAAACAAACCCGCTGGTAGGGTGGTTTT	3636
3668	QY	TTGTTTCAAGCAGCAGATTACCGCGCAGAAAAAAGATCTCAAGAAGATCTTTTGATCT	3727
3635	DB	TTGTTTCAAGCAGCAGATTACCGCGCAGAAAAAAGATCTCAAGAAGATCTTTTGATCT	3576
3728	QY	TTTTCTACGGGCTCGACGCTCAGTGGAAAGAAACCTCACTTAAAGGATTTTGGTCAAGA	3787
3575	DB	TTTTCTACGGGCTCGACGCTCAGTGGAAAGAAACCTCACTTAAAGGATTTTGGTCAAGA	3516
3788	QY	GATTATCGTCGACCAAGCGGCCATCGTGCT-----	3819
3515	DB	GATTATCAAAAAGGATTTTCACTTAGATCTTTTAAATTAAAAATGAAGTTTTAAATCAA	3456
3820	QY	-----CCCCCTCTCCAGTTCTCGGG	3840
3455	DB	TCATAAGTATATAGCTAACTCAGGCTATGGCAGGCGCTGCGCGCCGACGTTGGCTG	3396
3841	QY	GCATGGATGCGCGGATAGCGCTCTGGTTCTTGATGCGCAGGATTTGCATCGCGG	3900
3395	DB	CGAGCCCTGGGCTTCAACCCAACTTGGGGGGTGGGTGGGGAAGAAAGAACCGCGG	3336
3901	QY	-----TAGAACTCCGCGAGTGTGCTCCAGCTCAGGCGCAGCAGCTGAACCAAC	3946
3335	DB	CGATTGGCCCCAATGGGTCTCGTGGGGTATCGACAGAGTGCCAGCCCTGGGACCGAA	3276
3947	QY	TCGCGAGGGATCGA-----	3961
3275	DB	CCCCCGTTTATGAACAAACGACCCAAACCGCTGCTTTTATCTGCTTTTATTTGCGG	3216
3962	QY	-----GCCCGG	3967
3215	DB	TCATAGCGGGGTCTTCCGGTATTGTCTCTTCCGTGTTTTTCTCGGTAGCTCCCTTAG	3156
3968	QY	GGTGGCGGAAGAACTCCAGCATAGATCCCCGCGCTGGAGGATCATCCAGGGCGGCTCCC	4027

Db	3155	GGTGGGCGAAGAACTCCAGATCAGATCCCGCGCTGGAGATCATCCAGCCGCGCGTCCC	3096
Qy	4028	GGAAAAAGATTCGGAAGCCCAACTTTTCATAGAAAGCGGCGGTGGAAATCGAAATCTCGTG	4087
Db	3095	GGAAAAAGATTCGGAAGCCCAACTTTTCATAGAAAGCGGCGGTGGAAATCGAAATCTCGTG	3036
Qy	4088	ATGCGAGTTGGGCGTCCGCTTGCTGGTCAATTTGAAACCCACAGAGTCCCGCTCAGAAGAA	4147
Db	3035	ATGCGAGTTGGGCGTCCGCTTGCTGGTCAATTTGAAACCCACAGAGTCCCGCTCAGAAGAA	2976
Qy	4148	CTCGTCAAGAAGGCGATAGAAAGCGATGCGCTGCGAATCGGAGCGGCGGATACCGTAAAG	4207
Db	2975	CTCGTCAAGAAGGCGATAGAAAGCGATGCGCTGCGAATCGGAGCGGCGGATACCGTAAAG	2916
Qy	4208	CACAGGAAGCGGTACGCCCATTCGCGCCCAAGCTCTTCAGCAATATCATCGGTTAGCCAA	4267
Db	2915	CACAGGAAGCGGTACGCCCATTCGCGCCCAAGCTCTTCAGCAATATCATCGGTTAGCCAA	2856
Qy	4268	CGCTATGTCTGTATAGCGGTCGCGCCACACCCAGCGCGGCCACAGTTCGATGAATCCAGAAAA	4327
Db	2855	CGCTATGTCTGTATAGCGGTCGCGCCACACCCAGCGCGGCCACAGTTCGATGAATCCAGAAAA	2786
Qy	4328	CGGCGCAATTTTCCACCAATGATATTCGGGAAGCAGGCGATCGCCATGGCTCAGCAGAGATC	4387
Db	2795	CGGCGCAATTTTCCACCAATGATATTCGGGAAGCAGGCGATCGCCATGGCTCAGCAGAGATC	2736
Qy	4388	CTCGCGTTCGGGATCGCGGCTTTGAGCTGCGGAAACAGTTTCGGCTGGCGAGAGCCCGTG	4447
Db	2735	CTCGCGTTCGGGATCGCGGCTTTGAGCTGCGGAAACAGTTTCGGCTGGCGAGAGCCCGTG	2676
Qy	4448	ATGCTCTTTGTCAGAGATCATCTGATCGAACAGACCGGCTTCATTCGAGTACGTGCTCG	4507
Db	2675	ATGCTCTTTGTCAGAGATCATCTGATCGAACAGACCGGCTTCATTCGAGTACGTGCTCG	2618
Qy	4508	CTCGATGCGATGTTTCGCTTGGTTCGAATGGGCGAGTACCGCGATCAACGCGTATGCAG	4566
Db	2615	CTCGATGCGATGTTTCGCTTGGTTCGAATGGGCGAGTACCGCGATCAACGCGTATGCAG	2556
Qy	4568	CGCGCGCATTCATCAGCCATGATGGATACCTTTCTCGCAGGAGCAAGGTGAGATGACAG	4627
Db	2555	CGCGCGCATTCATCAGCCATGATGGATACCTTTCTCGCAGGAGCAAGGTGAGATGACAG	2496
Qy	4628	GAGATCTGTCGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGCACAC	4688
Db	2495	GAGATCTGTCGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGCACAC	2436
Qy	4688	GTGAGGACAGCTGCGCAAGGAAACGCCCGTCTGTGCCAGCCACGATAGCCGCGTGCCTC	4747
Db	2435	GTGAGGACAGCTGCGCAAGGAAACGCCCGTCTGTGCCAGCCACGATAGCCGCGTGCCTC	2376
Qy	4748	GTCCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGCGGCC	4807
Db	2375	GTCCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGCGGCC	2314
Qy	4808	CTGCGCTGACAGCCCGGAAACAGGCGGATCAGAGCAGCGGATGTCGTGTGTCGCCAGTTC	4866
Db	2315	CTGCGCTGACAGCCCGGAAACAGGCGGATCAGAGCAGCGGATGTCGTGTGTCGCCAGTTC	2255
Qy	4868	ATAGCCGAATAGCTCTCCACCCAAAGCGGCCCGGAGAACCTGCGTGCAATCCATCTGTTTC	4927
Db	2255	ATAGCCGAATAGCTCTCCACCCAAAGCGGCCCGGAGAACCTGCGTGCAATCCATCTGTTTC	2196
Qy	4928	AATCATGCGAAACGATTCCTCATCTGCTCTTTGATCAGATCTTTG	4971
Db	2195	AATCATGCGAAACGATTCCTCATCTGCTCTTTGATCAGATCTTTG	2152

RESULT 11
US-08-751-767A-7/C
; Sequence 7, Application US/08751767A
; Patent NO. 5994104
; GENERAL INFORMATION:

APPLICANT: ANDERSON, ROBERT J.
 APPLICANT: GRANT, HUGH
 APPLICANT: MACDONALD, IAN D.
 TITLE OF INVENTION: INTERLUKIN-12 FUSION PROTEIN
 NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHUYE P.C.
 STREET: 1100 NORTH GLSBE ROAD
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22201
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/08/751,767A
 APPLICATION NUMBER: US/08/751,767A
 FILING DATE: 08-NOV-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: SADOFF, B.J.
 REGISTRATION NUMBER: 36,663
 REFERENCE/DOCKET NUMBER: 117-221
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 7038164091
 TELEFAX: 7038164100
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6139 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3222..4841
 US-08-751-767A-7

Query Match 29.7%; Score 1570; DB 2; Length 6139;
 Best Local Similarity 86.2%; Pred. No. 0;
 Matches 1849; Conservative 0; Mismatches 115; Indels 180; Gaps 3;

QY	3008	ACTCAAGGCGGTAAATCGGTATCCACAGATCAGGGGATACGAGGAAACACATGT	3067
Db	2624	AATGCATGCGCGTAAATCGGTATCCACAGATCAGGGGATACGAGGAAACACATGT	2565
QY	3068	GAGCAAAAGGCGCAGCAAAAGGCGCAGGAAACCGTAAAGGCGCGTTCTGGCGTTTTC	3127
Db	2564	GAGCAAAAGGCGCAGCAAAAGGCGCAGGAAACCGTAAAGGCGCGTTCTGGCGTTTTC	2505
QY	3128	ATAGCTCCGCGCGCGCTGACGAGCATCACAAATTCACGCTCAAGTCAGAGTGGGAA	3187
Db	2504	ATAGCTCCGCGCGCGCTGACGAGCATCACAAATTCACGCTCAAGTCAGAGTGGGAA	2445
QY	3188	ACCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC	3247
Db	2444	ACCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC	2385
QY	3248	CTGTTCCGACCGCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGG	3307
Db	2384	CTGTTCCGACCGCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGG	2325
QY	3308	CGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTTCCAGC	3367
Db	2324	CGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTTCCAGC	2265
QY	3368	TGGGTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTAACTATC	3427
Db	2264	TGGGTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTAACTATC	2205
QY	3428	GTCCTTGAGTCCAAACCCGCTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA	3487

Db	2204	GTCTTGAGTCCAAACCCGCTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA	2145
QY	3488	GGATTACGAGCGAGCTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACT	3547
Db	2144	GGATTACGAGCGAGCTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACT	2085
QY	3548	ACGGCTACACTAGAAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAACCAAGTACCTTCG	3607
Db	2084	ACGGCTACACTAGAAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAACCAAGTACCTTCG	2025
QY	3608	GAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGAGCGGTGGTTTTT	3667
Db	2024	GAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGAGCGGTGGTTTTT	1965
QY	3668	TTGTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCTCTTTGATCT	3727
Db	1964	TTGTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCTCTTTGATCT	1905
QY	3728	TTTCTACGGGTCTGACGCTCAGTGGAAACGAAATCTCAGTTAAGGATTTTGGTCAAGA	3787
Db	1904	TTTCTACGGGTCTGACGCTCAGTGGAAACGAAATCTCAGTTAAGGATTTTGGTCAAGA	1845
QY	3788	GATTATCGTCGACCAAGCGCCATCGTGCCT	3819
Db	1844	GATTATCAAAAGGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAA	1785
QY	3820	-----CCCCACTCTGCAGTTTCGGGG	3840
Db	1784	TTCTAAGTATATATGATGAATACCTGAGGCTATGGCAGGCGCTCGCCGCCACGTTGGCTG	1725
QY	3841	GCATGGATGCGCGATAGCGCTGCTGTTTCTTGGATCCGACGCGATTTGCATCTGCGG	3900
Db	1724	GCAGCCCTGGGCTTACCCGAACTTGGGGGTGGGTGGGAAAGGAAGAACGCGGG	1665
QY	3901	-----TAGAACTCCGCGAGGTCCTCAGCTCAGCAGCAGCAGCTGAACCAAC	3946
Db	1664	CGTATTGCGCCCAATGGGTCTCGTGGGTATCGACAGAGTGCAGCCCTGGGACCGAA	1605
QY	3947	TCGCGAGGGGATCGA	3961
Db	1604	CCCCCGCTTATGAACAAACGACCAACACCGTCGCTTATTTCTGCTTTTATTTATTCGCG	1545
QY	3962	-----GCCGGG	3967
Db	1544	TCATAGCGGGTCTTCCGCTATTTCTCTTCCGTTTTCAGTTAGCTCCCTCCCTAG	1485
QY	3968	GGTGGCGAAGAACTCCAGCATGATATCCCGCGCTGGAGGATCATCCAGCGCGGCTCC	4027
Db	1484	GGTGGCGAAGAACTCCAGCATGATATCCCGCGCTGGAGGATCATCCAGCGCGGCTCC	1425
QY	4028	GGAAAAAGGATCCGAAAGCCCAACCTTTTATAGAAAGCGCGGTGGAATTCGAAATCTCGTG	4087
Db	1424	GGAAAAAGGATCCGAAAGCCCAACCTTTTATAGAAAGCGCGGTGGAATTCGAAATCTCGTG	1365
QY	4088	ATGCGAGTGGGCTGCTTGGTGGTCAATTTGAAACCCAGAGTCCCGCTCAGAGAA	4147
Db	1364	ATGCGAGTGGGCTGCTTGGTGGTCAATTTGAAACCCAGAGTCCCGCTCAGAGAA	1305
QY	4148	CTCGTCAGAAAGCGCATAGAGGCGATCGCTGGAATCGGAGCGCGATACCGTAAAG	4207
Db	1304	CTCGTCAGAAAGCGCATAGAGGCGATCGCTGGAATCGGAGCGCGATACCGTAAAG	1245
QY	4208	CACGAGGAGCGGTGAGCCCATTCGCCGCCCAAGCTCTTCAGCAATATCACGGGTAGCCAA	4267
Db	1244	CACGAGGAGCGGTGAGCCCATTCGCCGCCCAAGCTCTTCAGCAATATCACGGGTAGCCAA	1185
QY	4268	CGGTATGCTCATAGCGGTCCGACACACCGCGCCACAGTCAGTGAATTCAGAGAAA	4327
Db	1184	CGGTATGCTCATAGCGGTCCGACACACCGCGCCACAGTCAGTGAATTCAGAGAAA	1125
QY	4328	GGGGCCATTTCCACCATGATATTCGGAAGCAGGATCGCCCATGGGTCAAGCAGATC	4387

1124 GGGGCCATTTTCCACCATGATATTTCCGCAAGCAGGATGCCATGGGTACGACGAGATC 1065
QY 4388 CTCGGCTGGGCGATCGCGGCTTGGCTTGGCGCAACAGTTCCGCTGGCGGAGCCCTG 4447
Db 1064 CTCGGCTGGGCGATCGCTTGGCTTGGCGCAACAGTTCCGCTGGCGGAGCCCTG 1005
QY 4448 ATGCTCTTCTGCTCCAGATCATCTGATCGACAAAGCCGCTTCCATCCGAGTACGTGCTG 4507
Db 1004 ATGCTCTTCTGCTCCAGATCATCTGATCGACAAAGCCGCTTCCATCCGAGTACGTGCTG 945
QY 4508 CTCGATGCGATGTTTCTGCTGGTGGTGGATGGGCGAGTACCGATCAAGCGTATGCG 4567
Db 944 CTCGATGCGATGTTTCTGCTGGTGGTGGATGGGCGAGTACCGGATCAAGCGTATGCG 885
QY 4568 CGCGCGCATTCGATCAGCGCATGATGATATCTTCTCGCGAGGAGCAAGGTGAGATGACAG 4627
Db 884 CGCGCGCATTCGATCAGCGCATGATGATATCTTCTCGCGAGGAGCAAGGTGAGATGACAG 825
QY 4628 GAGATCTGCGCGCGCATTCGCGCAATAGCAGCGAGTCCCTTCCGCTTCAAGTGAAC 4687
Db 824 GAGATCTGCGCGCGCATTCGCGCAATAGCAGCGAGTCCCTTCCGCTTCAAGTGAAC 765
QY 4688 GTCGAGCAGCTGCGCAGGAAAGCCGCTGCTGCGCAGCGATAGCGCGCTGCTC 4747
Db 764 GTCGAGCAGCTGCGCAGGAAAGCCGCTGCTGCGCAGCGATAGCGCGCTGCTC 705
QY 4748 GTCCTGCGATTCATTCAGGCGACCGGACAGTCCGCTTGTGACAAAAGAACCGCGCGCC 4807
Db 704 GTCCTGCGATTCATTCAGGCGACCGGACAGTCCGCTTGTGACAAAAGAACCGCGCGCC 645
QY 4808 CTGCGCTGACGCGGACAGCGGCGCATCAGAGCGGATGCTGTTGTGCGCCAGTC 4867
Db 644 CTGCGCTGACGCGGAAACAGCGGCGCATCAGAGCGGATGCTGTTGTGCGCCAGTC 585
QY 4868 ATAGCGGAATAGCTCTCCACCAAGCGCGCGGAGAACCTCGTGCAATCCATCTTGTC 4927
Db 584 ATAGCGGAATAGCTCTCCACCAAGCGCGCGGAGAACCTCGTGCAATCCATCTTGTC 525
QY 4928 ATATGCGGAACGATCTCATCTGCTCTTGTATGATCATCTTG 4971
Db 524 ATATGCGGAACGATCTCATCTGCTCTTGTATGATCATCTTG 481

RESULT 12

US-08-723-306-6/c
Sequence 6, Application US/08723306
Patent No. 5856178
GENERAL INFORMATION:
APPLICANT: White PhD, Kenneth
APPLICANT: Morrey PhD, John
APPLICANT: Reed, William
TITLE OF INVENTION: Cassette for Expression of Lytic
TITLE OF INVENTION: Peptides in Mammalian Transgenic Organisms
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Trask Britt and Rossa
STREET: P.O. Box 2550
CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,306
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweigert PhD, Susan E
REGISTRATION NUMBER: 36,289

REFERENCE/DOCKET NUMBER: 2549
TELEPHONE: 8015321922
TELEFAX: 8015319168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8797 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Construct comprising Bos
DESCRIPTION: taurus beta casein 5' regulatory region plus genes encoding
DESCRIPTION: amphipathic peptide and green fluorescent protein"
HYPOTHETICAL: YES
ANTI-SENSE: NO
US-08-723-306-6

Query Match 29.7%; Score 1569.8; DB 2; Length 8797;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;

QY 3013 AAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGCAAGGAAACATGTGAGCA 3072
Db 8793 ATGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGCAAGGAAACATGTGAGCA 8734
QY 3073 AAGGCGCAAAAGCCAGGAAACCGTAAAGGCGGGTTCCTGCGCTTTTTCATAGG 3132
Db 8733 AAGGCGCAAAAGCCAGGAAACCGTAAAGGCGGGTTCCTGCGCTTTTTCATAGG 8674
QY 3133 CTCGCGCCCTTACGAGCATCACAATAATCGACGCTCAAGTCAGAGTGGGAAACCG 3192
Db 8673 CTCGCGCCCTTACGAGCATCACAATAATCGACGCTCAAGTCAGAGTGGGAAACCG 8614
QY 3193 ACAGCATATAAGATACAGCGGTTTCCCTCTGAAGCTCCCTGCTGCGCTCTCTGTT 3252
Db 8613 ACAGCATATAAGATACAGCGGTTTCCCTCTGAAGCTCCCTGCTGCGCTCTCTGTT 8554
QY 3253 CGGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTGGCGTT 3312
Db 8553 CGGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTGGCGTT 8494
QY 3313 TCTCATAGTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAAAGCTGGC 3372
Db 8493 TCTCATAGTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAAAGCTGGC 8434
QY 3373 TGTGTGACGAAACCCCGCTTACCGCGCTTACCGCGCTTATCGGTAACCTATCGTCT 3432
Db 8433 TGTGTGACGAAACCCCGCTTACCGCGCTTATCGGCGCTTATCGGTAACCTATCGTCT 8374
QY 3433 GAGTCCAAACCCCGGTAAAGACACGACTTATCGCACTGGCAGCGCACTGGTAAACGAT 3492
Db 8373 GAGTCCAAACCCCGGTAAAGACACGACTTATCGCACTGGCAGCGCACTGGTAAACGAT 8314
QY 3493 AGCAGAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTAACTACGCG 3552
Db 8313 AGCAGAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTAACTACGCG 8254
QY 3553 TACACTAGAAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCGAGTACCTTCGAAAA 3612
Db 8253 TACACTAGAAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCGAGTACCTTCGAAAA 8194
QY 3613 AGAGTGGTGTAGTCTTGTATCCCGGCAAAACAAACCGCTGTAGCGGTGTTTTTGT 3672
Db 8193 AGAGTGGTGTAGTCTTGTATCCCGGCAAAACAAACCGCTGTAGCGGTGTTTTTGT 8134
QY 3673 TGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCTCTTGTATCTTTCT 3732
Db 8133 TGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCTCTTGTATCTTTCT 8074
QY 3733 ACGGGGTCTCAGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGTTCATGATTA 3792
Db 8073 ACGGGGTCTCAGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGTTCATGATTA 8014

3793 TGTGACCAAGCGCCATCGTGCT 3819
Db TCAGAAAGGATCTTACCTAGATCTTTTAAATTAAGATGAAGTTTAAATCAATCTAA 7954
3820 -----CCCCACTCTGAGTTTGGGGGATG 3845
Db AGTATATAGTAACCTGAGGCTATGGCAGGCTTGGCCCGACGTTGGCTCGAGC 7894
3846 GATGCGCGGATAGCCCTGCTGCTTCTTCTGATGCGACGATTTGCACTGCGG 3900
Db CTGGGCTTCAACCGAATTTGGGGGTGGGTGGGAAAGGAAGAACGCGGCGTAT 7834
3901 -----TAGAATTCGCGAGGTGCTCAGCCTCAGCAGAGCTGAACCAATCTCGG 3951
Db TGGCCCAATGGGCTCTCGTGGGTATCGACAGAGTGCACGCTTGGGACCGAACCCG 7774
3952 AGGGATCGA 3961
Db CGTTATGAACAAAGACCAACACCGTGGTATTTATCTGCTTTTATTCGCGTATA 7714
3962 -----GCCCGGGTGG 3972
Db GCGCGGTTCTTCGGTATGTCCTCTCGGTGTTTCACTAGCTCCCTAGGGTGG 7654
3973 GCGAAGAACTCCAGCATGATCCCGCGCTGGAGATCATCCAGCGCGCTCCGAAA 4032
Db GCGAAGAACTCCAGCATGATCCCGCGCTGGAGATCATCCAGCGCGCTCCGAAA 7594
4033 ACATTTCCGAAGCCCACTTTTATAGAGGGCGGTGGAATCGATGATGCG 4092
Db ACATTTCCGAAGCCCACTTTTATAGAGGGCGGTGGAATCGATGATGCG 7534
4093 AGTTGGGCTGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGTCAAGAACTCGT 4152
Db AGTTGGGCTGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGTCAAGAACTCGT 7474
4153 CRAAGAGGATAGAGGCGATCGCTGCGAATCGGGAGCGGATACCGTAAAGCAGA 4212
Db CAAGAGGCGATAGAGGCGATCGCTGCGAATCGGGAGCGGATACCGTAAAGCAGA 7414
4213 GGAAGCGGTGAGCCCAATTCGCGCAAGTCTTTCAGCAATATCACGGTAGCAACGCTA 4272
Db GGAAGCGGTGAGCCCAATTCGCGCAAGTCTTTCAGCAATATCACGGTAGCAACGCTA 7354
4273 TGTCTGTAGAGCGGTGGGACACACCGGCGGCGAGTGCATGATCCAGAAAGCGG 4332
Db TGTCTGTAGAGCGGTGGGACACACCGGCGGCGAGTGCATGATCCAGAAAGCGG 7294
4333 CATTTTCCACATGATATTCGCAAGCAGGATCGCCATGGGTACGACGAGATCTCGC 4392
Db CATTTTCCACATGATATTCGCAAGCAGGATCGCCATGGGTACGACGAGATCTCGC 7234
4393 GTGCGGCGATCGCGCTTGAAGCTGCGAAGTTCGCTGGCGGCGAGCCCTGATGCT 4452
Db GTGCGGCGATCGCGCTTGAAGCTGCGAAGTTCGCTGGCGGCGAGCCCTGATGCT 7174
4453 CTTCGTCCAGATCATCTGATCGACAGACCGGCTTCCATCCAGTACGCTCGCTCGA 4512
Db CTTCGTCCAGATCATCTGATCGACAGACCGGCTTCCATCCAGTACGCTCGCTCGA 7114
4513 TGGATTTTTCGTTGGTGGTGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGC 4572
Db TGGATTTTTCGTTGGTGGTGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGC 7054
4573 GCATTGATCAGCATGATGATGATCTTCTCGCAGGAGCAAGGTGAGATGACAGAGAT 4632
Db GCATTGATCAGCATGATGATGATCTTCTCGCAGGAGCAAGGTGAGATGACAGAGAT 6994
4633 CTTGCCCGGCACTTCGCCCAATAGCAGCGAGTCCCTTCCGCTTCAAGTGAACGCTCGA 4692
Db CTTGCCCGGCACTTCGCCCAATAGCAGCGAGTCCCTTCCGCTTCAAGTGAACGCTCGA 6934

4693 GCACAGTGGCAGGAAGCCCGTCTGTGGCCAGCAGATAGCCGCTGCTCGTCT 4752
Db GCACAGTGGCAGGAAGCCCGTCTGTGGCCAGCAGATAGCCGCTGCTCGTCT 6874
4753 GCAGTTCAATTCAGGCGACCGGACAGGTCGCTTTGACAAAAGAACCGGCGCCCTGCG 4812
Db GCAGTTCAATTCAGGCGACCGGACAGGTCGCTTTGACAAAAGAACCGGCGCCCTGCG 6814
4813 CTGACAGCCGGAACACGCGGCGATCAGAGCAGCGATGCTGTTGTCGCTGATGAGC 4872
Db CTGACAGCCGGAACACGCGGCGATCAGAGCAGCGATGCTGTTGTCGCTGATGAGC 6754
4873 CGATAGCTCTTCAACCGAGCGCGGAGAACCTCGGTGCAATCCATCTTGTTCATCA 4932
Db CGATAGCTCTTCAACCGAGCGCGGAGAACCTCGGTGCAATCCATCTTGTTCATCA 6694
4933 TGGAAACGATCTCATCTCTCTCTTGTATCAGATCTTG 4971
Db TGGAAACGATCTCATCTCTCTCTTGTATCAGATCTTG 6655

RESULT 13

PCT-US96-10041-6/c
; Sequence 6, Application PC/TUS9610041
; GENERAL INFORMATION:
; APPLICANT: White Phd, Kenneth
; APPLICANT: Morrey Phd, John
; APPLICANT: Reed, William
; TITLE OF INVENTION: Cassette for Expression of Lytic
; TITLE OF INVENTION: Peptides in Mammalian Transgenic Organisms
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10041
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Phd, Susan E
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 8015321922
; TELEFAX: 8015319168
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8797 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Construct comprising Bos
; DESCRIPTION: taurus beta casein 5' regulatory region plus genes encoding
; DESCRIPTION: amphipathic peptide and green fluorescent protein"
; HYPOTHEtical: YES
; ANTI-SENSE: NO

PCT-US96-10041-6

Query Match 29.7%; Score 1569.8; DB 5; Length 8797;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;

QY 3013 AAGCGGTAAATACGGTTATCCACAGATCAAGGGGATAACGAGGAAACATGTGAGCA 3072

Db 8793 ATGCGGTAATACGGTATATCAAGATCAGGGATACACAGAAAGAACTGTGAGCA 8734
Qy 3073 AAAGGCGAGCAAAAGGCGAGCAACCGTAAAGAGCGCGGTGTGTGCGTTTTCATAGG 3132
Db 8733 AAAGGCGAGCAAAAGGCGAGCAACCGTAAAGAGCGCGGTGTGTGCGTTTTCATAGG 8674
Qy 3133 CTCGCGCCCTGACGAGCATCAGAAATCGAGCTCAAGTCAGAGGTGGCGAAACCG 3192
Db 8673 CTCGCGCCCTGACGAGCATCAGAAATCGAGCTCAAGTCAGAGGTGGCGAAACCG 8614
Qy 3193 ACAGGACTATAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTGTT 3252
Db 8613 ACAGGACTATAAGATACAGGGGTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTGTT 8554
Qy 3253 CCGACCTGCGGTTACCGGATACCTGTCGCTTCTCCCTTCCGGAAGCGTGGCGCTT 3312
Db 8553 CCGACCTGCGGTTACCGGATACCTGTCGCTTCTCCCTTCCGGAAGCGTGGCGCTT 8494
Qy 3313 TCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTGTAGTCTGCTCCAAAGCTGGG 3372
Db 8493 TCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTGTAGTCTGCTCCAAAGCTGGG 8434
Qy 3373 TGTGTGACGAAACCCCGGTTACCGGATACCTGTCGCTTCTCCCTTCCGGAAGCGTGGCGCTT 3432
Db 8433 TGTGTGACGAAACCCCGGTTACCGGATACCTGTCGCTTCTCCCTTCCGGAAGCGTGGCGCTT 8374
Qy 3433 GAGTCCAAACCGGTAAGACAGCATCTATCGCACTGGCAGCAGCACTGGTAACAGGATT 3492
Db 8373 GAGTCCAAACCGGTAAGACAGCATCTATCGCACTGGCAGCAGCACTGGTAACAGGATT 8314
Qy 3493 AGCAGAGGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTACGCG 3552
Db 8313 AGCAGAGGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTACGCG 8254
Qy 3553 TACACTAGAAAGAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACTTCGGAATA 3612
Db 8253 TACACTAGAAAGAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACTTCGGAATA 8194
Qy 3613 AGAGTGTAGTCTTTGATTCGCGCAAAACCAACCGTGTAGCGGTGTTTTTTTGT 3672
Db 8193 AGAGTGTAGTCTTTGATTCGCGCAAAACCAACCGTGTAGCGGTGTTTTTTTGT 8134
Qy 3673 TGCAGAGCAGATTTACGCGCAGAAAGAGTCTCAAGAGATCTTTGATCTTTTCT 3732
Db 8133 TGCAGAGCAGATTTACGCGCAGAAAGAGTCTCAAGAGATCTTTGATCTTTTCT 8074
Qy 3733 ACGGGTCTGACGCTCAGTGAACGAAACTCAGTTAAGGATTTTGGTCAATGAGATTA 3792
Db 8073 ACGGGTCTGACGCTCAGTGAACGAAACTCAGTTAAGGATTTTGGTCAATGAGATTA 8014
Qy 3793 TCGTGCACCAAGCGGCTATCGTCTCTTTAAATTAAGTCTTAAATCAATCTAA 3819
Db 8013 TCAAAAGGATCTTCACTAGATCTTTTAAATTAAGTCTTAAATCAATCTAA 7954
Qy 3820 -----CCCCACTCTCAGTTCGGGGGATG 3845
Db 7953 AGTATATAGTAACCTAGGCTATGCGAGGCTGCGCCCGAGCGTTGGCTGCGAG 7894
Qy 3846 GATGCGGATAGCGCTGCTGTTTCTGATGCGGAGGATTTGACTGCGG----- 3900
Db 7893 CTTGGGCTTACCCGAACTTTGGGGGTGGGGTGGGAAAGGAAACCGGGCGTAT 7834
Qy 3901 -----TAGAACTCGCGAGGTGCTCCAGCTCAGGACAGCTGAACCAACTGCG 3951
Db 7833 TGGCCCAATGGGTCTCGTGGGTATCGACAGTGCAGCCCTGGGACCGAACCCTG 7774
Qy 3952 AGGGATCA----- 3961
Db 7773 CGTTTATGAACAAAGACCAACCGTGGTTTTTATCTGTCTTTTATTTGCGGTCA 7714
Qy 3962 -----GCCCGGGTGG 3972

Db 7713 GCGGGGTCTCTCCGGTATTGTCTCTCGTGTCTTTCAGTTAGCTCTCCCTAGGCTGG 7654
Qy 3973 CCGAAGAACTCCAGATGAGATCCCGGCTGGAGGATCATCCAGCGGGTCCCGGAA 4032
Db 7653 CCGAAGAACTCCAGATGAGATCCCGGCTGGAGGATCATCCAGCGGGTCCCGGAA 7594
Qy 4033 ACGATTCCGAAGCCCAACCTTTTATAGAAGCGCGGTGGAATCGAAATCTCGTGTGGC 4092
Db 7593 ACGATTCCGAAGCCCAACCTTTTATAGAAGCGCGGTGGAATCGAAATCTCGTGTGGC 7534
Qy 4093 AGTTTGGGCTCGCTTGGTGGTCAATTTGNAACCCAGAGTCCCGCTCAGAAAGTCTGT 4152
Db 7533 AGTTTGGGCTCGCTTGGTGGTCAATTTGNAACCCAGAGTCCCGCTCAGAAAGTCTGT 7474
Qy 4153 CAAGAAGCGGATAGAAGCGGATGCGCTCGGAATCGGAGCGGCGATACCGTAAAGCAGA 4212
Db 7473 CAAGAAGCGGATAGAAGCGGATGCGCTCGGAATCGGAGCGGCGATACCGTAAAGCAGA 7414
Qy 4213 GGAAGCGGTGAGCGGATTTGCGCGCAAGCTCTTTCAGCAATATCAGGATGAGCGGAGCGTA 4272
Db 7413 GGAAGCGGTGAGCGGATTTGCGCGCAAGCTCTTTCAGCAATATCAGGATGAGCGGAGCGTA 7354
Qy 4273 TGTCTGTATAGCGGTCCGCCACACCCAGCGCGCCACAGTCCGATGAATCCAGAAAGCGGC 4332
Db 7353 TGTCTGTATAGCGGTCCGCCACACCCAGCGCGCCACAGTCCGATGAATCCAGAAAGCGGC 7294
Qy 4333 CATTTTCCACATGATATTCGGCAAGCGGATTCGCCATGCGTCAAGAGATGCTCGC 4392
Db 7293 CATTTTCCACATGATATTCGGCAAGCGGATTCGCCATGCGTCAAGAGATGCTCGC 7234
Qy 4393 CQTGGGATGCGGCTTGGCTGCGAAGCAGTTTGGCTGGGCGGAGCGGCTGATGCT 4452
Db 7233 CQTGGGATGCTCGCTTGGCTGCGAAGCAGTTTGGCTGGGCGGAGCGGCTGATGCT 7174
Qy 4453 CTTCTGCAGATCATCTGATCGAAGCAGCGGTTTCATCCGAGTACGTCGCTGCTGA 4512
Db 7173 CTTCTGCAGATCATCTGATCGAAGCAGCGGTTTCATCCGAGTACGTCGCTGCTGA 7114
Qy 4513 TCGGATGTTTGGCTTGGTGGTGAATGGCAGGTAGCGGATCAAGCGTATGACGCGCC 4572
Db 7113 TCGGATGTTTGGCTTGGTGGTGAATGGCAGGTAGCGGATCAAGCGTATGACGCGCC 7054
Qy 4573 GCATTGCTACGCTATGATGATCTTCTCGGAGGAGCAAGGTGAGTACAGGAGAT 4632
Db 7053 GCATTGCTACGCTATGATGATCTTCTCGGAGGAGCAAGGTGAGTACAGGAGAT 6994
Qy 4633 CTTGCGCGGCTTTCGCGCATAGCAGCGCTTCCCGCTTCCGCTTCACTGACGCTGA 4692
Db 6993 CTTGCGCGGCTTTCGCGCATAGCAGCGCTTCCCGCTTCCGCTTCACTGACGCTGA 6934
Qy 4693 GCACAGCTGCGCAAGGAAACCGCTGCTGGCCAGCAGTACGCGCTGCTGCTGCT 4752
Db 6933 GCACAGCTGCGCAAGGAAACCGCTGCTGGCCAGCAGTACGCGCTGCTGCTGCT 6874
Qy 4753 GAGTTTATTCAGGCAACCGGACGCTGCTTTCAGAAAGACCGGCGCGCTCGG 4812
Db 6873 GAGTTTATTCAGGCAACCGGACGCTGCTTTCAGAAAGACCGGCGCGCTCGG 6814
Qy 4813 CTGACAGCGGAAACACCGCGCATCAGAGCAGCGGATTTGTCTGTTGTGCCAGTATAGC 4872
Db 6813 CTGACAGCGGAAACACCGCGCATCAGAGCAGCGGATTTGTCTGTTGTGCCAGTATAGC 6754
Qy 4873 CGAATAGCTCTCCACCGGAGCGGAGAACCTTGGTGCATCCATCTTGTTCATCA 4932
Db 6753 CGAATAGCTCTCCACCGGAGCGGAGAACCTTGGTGCATCCATCTTGTTCATCA 6694
Qy 4933 TCGAAACCATCTCATCTCTCTTGTATCAGATCTTG 4971
Db 6693 TCGAAACCATCTCTCATCTCTCTTGTATCAGATCTTG 6655

RESULT 14

US-08-723-306-5/c

Sequence 5, Application US/08723306
Patent No. 5856178
GENERAL INFORMATION:
APPLICANT: White PhD, Kenneth
APPLICANT: Morrey PhD, John
APPLICANT: Read, William
TITLE OF INVENTION: Cassette for Expression of Lytic
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Trask Britt and Rossa
STREET: P.O. Box 2550
CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,306
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweigert PhD, Susan E
REGISTRATION NUMBER: 36,289
REFERENCE/DOCKET NUMBER: 2549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 8015321922
TELEFAX: 8015319168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 11093 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Construct comprising
portions of Bos taurus beta casein gene and genes encoding
amphipathic peptide and green fluorescent protein"
HYPOTHETICAL: YES
ANTI-SENSE: NO
FEATURE:
NAME/KEY: exon
LOCATION: 1801..1834
OTHER INFORMATION: /product= "beta casein exon 1"
FEATURE:
NAME/KEY: exon
LOCATION: 3780..3832
OTHER INFORMATION: /product= "beta casein exon 2"
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1766..1773
FEATURE:
NAME/KEY: exon
LOCATION: 4567..4590
OTHER INFORMATION: /product= "portion of beta casein
exon 3"
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 127..1800
OTHER INFORMATION: /function= "5' flanking regulatory
region of bovine beta casein gene"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 4587..5310
OTHER INFORMATION: /product= "Green fluorescent
protein"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 5320..5449

OTHER INFORMATION: /product= "Shiva-1 coding sequence"
FEATURE:
NAME/KEY: polyA_signal
LOCATION: 7630..7635
OTHER INFORMATION: /standard_name= "Bovine beta casein"
OTHER INFORMATION: 3' region, in exon 9"
US-08-723-306-5
Query Match 29.7%; Score 1569.8; DB 2; Length 11093;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;
QY 3013 AAGCGGTATACGGTTATCCACAGATCAGGGGATAACCGCAGGAAGAACATGTGAGCA 3072
DB 11089 ATGGCGGTATACGGTTATCCACAGATCAGGGGATAACCGCAGGAAGAACATGTGAGCA 11030
QY 3073 AAAGGCCAGCAAAAGGCCAGAACCCGTAATAAGGCGCGTTGCTGGCGTTTTTCCATAGG 3132
DB 11029 AAAGGCCAGCAAAAGGCCAGAACCCGTAATAAGGCGCGTTGCTGGCGTTTTTCCATAGG 10970
QY 3133 CTCGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG 3192
DB 10969 CTCGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG 10910
QY 3193 ACAGGACTATAAGATACACAGGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTCTTT 3252
DB 10909 ACAGGACTATAAGATACACAGGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTCTTT 10850
QY 3253 CCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTGGGAAGCGTGGCGCTT 3312
DB 10849 CCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTGGGAAGCGTGGCGCTT 10790
QY 3313 TCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTTCGCTCAAGCTGGGC 3372
DB 10789 TCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTTCGCTCAAGCTGGGC 10730
QY 3373 TGTGTGACGAACCCCGCTTACGCGCCGACCGCTGGCGCTTATCCCGTAATCTATCGTCTT 3432
DB 10729 TGTGTGACGAACCCCGCTTACGCGCCGACCGCTGGCGCTTATCCCGTAATCTATCGTCTT 10670
QY 3433 GAGTCCAAACCGGTAAAGACAGCATTCGCCACTGGCAGCAGCCACTGTGTACAGATT 3492
DB 10669 GAGTCCAAACCGGTAAAGACAGCATTCGCCACTGGCAGCAGCCACTGTGTACAGATT 10610
QY 3493 AGCAGAGCGAGGTATGTAGCGGTGTACAGAGTCTCTGAAGTGTGGCTTAACACTACGC 3552
DB 10609 AGCAGAGCGAGGTATGTAGCGGTGTGTACAGAGTCTCTGAAGTGTGGCTTAACACTACGC 10550
QY 3553 TACACTAGAAGAACAGTATTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGAAAA 3612
DB 10549 TACACTAGAAGAACAGTATTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGAAAA 10490
QY 3613 AGAGTTGGTAGCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTTTGT 3672
DB 10489 AGAGTTGGTAGCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTTTGT 10430
QY 3673 TGCAAGCAGCAGATTACGCGCAGAAAAAAGAGTCTCAAGAGATCCTTTGATCTTTCT 3732
DB 10429 TGCAAGCAGCAGATTACGCGCAGAAAAAAGAGTCTCAAGAGATCCTTTGATCTTTCT 10370
QY 3733 ACGGGGTCTACGCTCAGTGGAAACGAAAACTCAGTTAAGGGAATTTGTCATGAGATTA 3792
DB 10369 ACGGGGTCTACGCTCAGTGGAAACGAAAACTCAGTTAAGGGAATTTGTCATGAGATTA 10310
QY 3793 TCGTCGACCAAGCGGCACTCGTGCCT----- 3819
DB 10309 TCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAGTATTAAATCAATCTAA 10250
QY 3820 -----CCCCACTCTCTGAGTTTCGGGGGATG 3845
DB 10249 AGTATATATGAGTAACTGAGGCTATGCGAGGCGCTGCCCGCCGACGTTGGCTGGAGC 10190
QY 3846 GATCGCGGATAGCGCGTGTGCTTCTTGATGCCGAGATTGCACTGCCG----- 3900

Db 10189 CCTGGCCTTACCCGAACTTGGGGGTGGGTGGGAAGAGAAACGCGGGCGTAT 10130
QY 3901 -----TGAAGACTCGGAGAGTGTCTCAGCCTCAGGACAGAGCTGAACCAACTCGCG 3951
Db 10129 TGGCCCCAATGGGTCTCGGTGGGTATCGACAGAGTGCAGCCCTGGGACCAACCCCG 10070
QY 3952 AGGGGATCGA----- 3961
Db 10069 GGTATGAACAACAGCCCAACACCGTGGGTATTTCTGTCTTTTATTGCGGTGCTATA 10010
QY 3962 -----GCCCGGGGTGG 3972
Db 10009 GCGGGGTTCCTTCGGTATTGTCTCCTTCGGTGTTCAGTTAGCTTCCCTCCTAGGCTGG 9950
QY 3973 GCGAAGAACTCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGGTCCCGGAAA 4032
Db 9949 GCGAAGAACTCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGGTCCCGGAAA 9890
QY 4033 ACGATTCCGAAGCCCAACCTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTATGCG 4092
Db 9889 ACGATTCCGAAGCCCAACCTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTATGCG 9830
QY 4093 AGGTGGCGCTGCTGTGGTGGTCTATTCGAACCCAGAGTCCCGCTCAGAGAACTCGT 4152
Db 9829 AGGTGGCGCTGCTGTGGTGGTCTATTCGAACCCAGAGTCCCGCTCAGAGAACTCGT 9770
QY 4153 CAAGAAGCGCATAGAAGCGCATGGCTGCGAATCGGAGCGGGGATACCGTAAGACACGA 4212
Db 9769 CAAGAAGCGCATAGAAGCGCATGGCTGCGAATCGGAGCGGGGATACCGTAAGACACGA 9710
QY 4213 GGAAGCGCTCAGCCATTCGCGCGCAAGCTTTCAGCAATATCAGCGGTAGCCACGCTA 4272
Db 9709 GGAAGCGCTCAGCCATTCGCGCGCAAGCTTTCAGCAATATCAGCGGTAGCCACGCTA 9650
QY 4273 TGCTCTGATAGCGTCCGCCACACCCAGCGCGGCACACAGTGCATGAATCCAGAAAAGCGC 4332
Db 9649 TGCTCTGATAGCGTCCGCCACACCCAGCGCGGCACACAGTGCATGAATCCAGAAAAGCGC 9590
QY 4333 CATTTTCCACATGATATTCGGGAGCGAGGATGCCATGGGTACAGCAGAGATCTTCGC 4392
Db 9589 CATTTTCCACATGATATTCGGGAGCGAGGATGCCATGGGTACAGCAGAGATCTTCGC 9530
QY 4393 CGTCGGGATCGCGCTTGAGCTGGCGAACAGTTCGGCTGGCGGAGCGCCCTGATGCT 4452
Db 9529 CGTCGGGATCGCTGCTTGAGCTGGCGAACAGTTCGGCTGGCGGAGCGCCCTGATGCT 9470
QY 4453 CTTCGTCAGATCATCTCTGATCGAACAAGACCGGCTTCATCCGAGTACGTGCTCGTCTGA 4512
Db 9469 CTTCGTCAGATCATCTCTGATCGAACAAGACCGGCTTCATCCGAGTACGTGCTCGTCTGA 9410
QY 4513 TGCGATGTTTCGCTTGGTGGTTCGAATGGGAGGTAGCGGATCAAGCGTATGACGCGCC 4572
Db 9409 TGCGATGTTTCGCTTGGTGGTTCGAATGGGAGGTAGCGGATCAAGCGTATGACGCGCC 9350
QY 4573 GCATTGATCAGCCATGATGATATCTTCTCGGAGGAGCAAGGTGAGATGACAGGAGAT 4632
Db 9349 GCATTGATCAGCCATGATGATATCTTCTCGGAGGAGCAAGGTGAGATGACAGGAGAT 9290
QY 4633 CTGCGCGCGGACTTCGCGCCATAGCAGCCAGTCCCTTCCCGCTTCAGTGACACGTCGA 4592
Db 9289 CTGCGCGCGGACTTCGCGCCATAGCAGCCAGTCCCTTCCCGCTTCAGTGACACGTCGA 9230
QY 4693 GCACAGCTCGCAGAGAAACCGCGCTGCTGGCGCAGCCAGATAGCGCGCTGCTCGCTCT 4752
Db 9229 GCACAGCTCGCAGAGAAACCGCGCTGCTGGCGCAGCCAGATAGCGCGCTGCTCGCTCT 9170
QY 4753 GCAGTTTCAATCAGGCGACCGGACAGGTGGTCTTGAACAAAGAACCGGGGCGCCCTGCG 4812
Db 9169 GCAGTTTCAATCAGGCGACCGGACAGGTGGTCTTGAACAAAGAACCGGGGCGCCCTGCG 9110
QY 4813 CTGACAGCGCGAACAACGGCGGATCAGAGCAGCGGATGCTGTGTCGCCAGTCATAGC 4872

RESULT 15

PCT-US96-10041-5/c
; Sequence 5, Application PC/TUS9610041
; GENERAL INFORMATION:
; APPLICANT: White PhD, Kenneth
; APPLICANT: Morrey PhD, John
; APPLICANT: Reed, William
; TITLE OF INVENTION: Cassette for Expression of Lytic
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10041
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert PhD, Susan E
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 8015321922
; TELEFAX: 8015319168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11093 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Construct comprising
; DESCRIPTION: portions of Bos taurus beta casein gene and genes encoding
; DESCRIPTION: amphipathic peptide and green fluorescent protein"
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1801..1834
; OTHER INFORMATION: /product= "beta casein exon 1"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3780..3832
; OTHER INFORMATION: /product= "beta casein exon 2"
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1766..1773
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4567..4590
; OTHER INFORMATION: /product= "portion of beta casein
; FEATURE: exon 3"

/ NAME/KEY: 5'UTR
/ LOCATION: 127..1800
/ OTHER INFORMATION: /function= "5' flanking regulatory
/ OTHER INFORMATION: region of bovine beta casein gene"
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 4587..5310
/ OTHER INFORMATION: /product= "Green fluorescent
/ OTHER INFORMATION: protein"
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 5320..5449
/ OTHER INFORMATION: /product= "Shiva-1 coding sequence"
/ FEATURE:
/ NAME/KEY: polyA_signal
/ LOCATION: 7630..7635
/ OTHER INFORMATION: /standard name= "Bovine beta casein
/ OTHER INFORMATION: 3' region, in exon 9"
PCT-US96-10041-5

Query Match 29.7%; Score 1569.8; DB 5; Length 11093;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;
QY 3013 AAGGCGGTATACGGTTATCCAGAAATCAGGGGATAACGAGGAAAGAAATGTGAGCA 3072
DB 11089 ATGGCGGTATACGGTTATCCACAGATCAGGGGATTAACGAGGAAAGAAATGTGAGCA 11030
QY 3073 AAAGGCCAGCAAAAGCCAGGAACCGTAAAGGCGCGTTGCTGCGCTTTTCCATAGG 3132
DB 11029 AAAGGCCAGCAAAAGCCAGGAACCGTAAAGGCGCGTTGCTGCGCTTTTCCATAGG 10970
QY 3133 CTCGCGCCCTCAGCAGCATCACAAAATCAGCCTCAGTCAAGTGGCGTCTCCTGTT 3192
DB 10969 CTCGCGCCCTCAGCAGCATCACAAAATCAGCCTCAGTCAAGTGGCGTCTCCTGTT 10910
QY 3193 ACAGGACTATAAGTATACAGGCGTTTCCCTCGGAAGTCCCTCGTGGCTCTCCTGTT 3252
DB 10909 ACAGGACTATAAGTATACAGGCGTTTCCCTCGGAAGTCCCTCGTGGCTCTCCTGTT 10850
QY 3253 CGACCCCTCGGTTACCGGATACCTGCGGCTTTTCCCTCGGGAAGCGTGGCGTT 3312
DB 10849 CGACCCCTCGGTTACCGGATACCTGCGGCTTTTCCCTCGGGAAGCGTGGCGTT 10790
QY 3313 TCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTCTCGTCCAAAGTGGGC 3372
DB 10789 TCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTCTCGTCCAAAGTGGGC 10730
QY 3373 TGTGTGCAAGACCCCGGTTAGCGCCGACCGCTGCGCTTATCCGGTAACTATGCTTT 3432
DB 10729 TGTGTGCAAGACCCCGGTTAGCGCCGACCGCTGCGCTTATCCGGTAACTATGCTTT 10670
QY 3433 GAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCACTGGTAAAGGATT 3492
DB 10669 GAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCACTGGTAAAGGATT 10610
QY 3493 AGCAGAGCAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGCG 3552
DB 10609 AGCAGAGCAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGCG 10550
QY 3553 TACACTAGAAGACAGTATTTGGTATCTGCGTCTGCTGGAAGCCAGTTACCTTCGGAATA 3612
DB 10549 TACACTAGAAGACAGTATTTGGTATCTGCGTCTGCTGGAAGCCAGTTACCTTCGGAATA 10490
QY 3613 AGAGTTGGTAGTCTTTGATCCCGCAAAACAAACACCGTGGTGGTGGTGGTGGTGGTGGT 3672
DB 10489 AGAGTTGGTAGTCTTTGATCCCGCAAAACAAACACCGTGGTGGTGGTGGTGGTGGTGGT 10430
QY 3673 TGCAAGCAGCATTTACGCGCAGAAAAGGATCTCAAGAGATCTTTGATCTTTTCT 3732
DB 10429 TGCAAGCAGCATTTACGCGCAGAAAAGGATCTCAAGAGATCTTTGATCTTTTCT 10370
QY 3733 ACGGGGTCTGACGCTCAGTGGTGAACGAAACTCAGGTTAAGGGATTTTGGTCAATGAGATTA 3792

DB 10369 ACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTGGTCAATGAGATTA 10310
QY 3793 TCGTCCGACCAAGCGGCCATCGTCCCT----- 3819
DB 10309 TCAAAAAGGATCTTACCTAGATCTCTTTTAAATATAAAGTATTAATCAATCTAA 10250
QY 3820 -----CCCCACTCTCTGAGTTTCGGGGCATG 3845
DB 10249 AGTATATATGATTAACCTGAGGCTATGCGAGGCGCTGCGCCGACGTTGGTGGCAGC 10190
QY 3846 GATCGCGGATAGCGCTGCTGCTTCTTGGATCCGACGATTTGCACTGCCGG----- 3900
DB 10189 CTTGGGCTTTCACCCGAACCTTGGGGGTGGGGTGGGAAAGAAAGAACGCGGGGTAT 10130
QY 3901 -----TAGAACTCCGCGAGTCTGTCAGCCTCAGGCGAGCTGAACCAACTCGCG 3951
DB 10129 TGGCCCCAATGGGCTCTCGTGGGGTATCGACAGATGTCAGCCCTGGGACCGAACCCG 10070
QY 3952 AGGGATCGA----- 3961
DB 10069 CGTTTATGAACAAACGACCAACACCGTGGCTTTTATTCTGCTTTTATTGCGCTATA 10010
QY 3962 -----GCCCGGGTGG 3972
DB 10009 GCGGGGTCTCTCCGGTATTGCTCTCTTCCGTGTTTTCAGTTAGCCTCCCCCTAGGGTGG 9950
QY 3973 GCAGAAACTCCAGCATGATCCCCCGCTGGAGATCATCAGCGCGGCTCCCGGAA 4032
DB 9949 GCAGAAACTCCAGCATGATCCCCCGCTGGAGATCATCAGCGCGGCTCCCGGAA 9990
QY 4033 ACATTTCCGAAGCCCACTTTTCATAGAGGCGGCTGGAATCGAAATCTCTGTATGGC 4092
DB 9889 ACATTTCCGAAGCCCACTTTTCATAGAGGCGGCTGGAATCGAAATCTCTGTATGGC 9930
QY 4093 AGGTGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4152
DB 9829 AGGTGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9770
QY 4153 CAAGAGCGCATAGAGGCGATCGCTGCGATCGGAGGCGGATACCGTAAAGCAGCA 4212
DB 9769 CAAGAGCGCATAGAGGCGATCGCTGCGATCGGAGGCGGATACCGTAAAGCAGCA 9710
QY 4213 GGAAGCGGTGAGCCCTTCCCGCCCAAGCTCTTCAGCAATATTCACGGGTAGCCAAAGCCTA 4272
DB 9709 GGAAGCGGTGAGCCCTTCCCGCCCAAGCTCTTCAGCAATATTCACGGGTAGCCAAAGCCTA 9650
QY 4273 TGTCTGTATAGCGCTGCGCCACACCGAGCGGCGCACTGCTGATGATTCAGAAAGCGGC 4332
DB 9649 TGTCTGTATAGCGCTGCGCCACACCGAGCGGCGCACTGCTGATGATTCAGAAAGCGGC 9590
QY 4333 CATTTTCCACCATGATATTCGGCAAGCAGGCTGCGCATGCGGTCAGAGATCTCTCGC 4392
DB 9589 CATTTTCCACCATGATATTCGGCAAGCAGGCTGCGCATGCGGTCAGAGATCTCTCGC 9530
QY 4393 CGTGGGCACTGCGCGCTTGAAGCTGCGCAAGCTTGGCTGCGGAGCGGCTGCTGATGCT 4452
DB 9529 CGTGGGCACTGCGCGCTTGAAGCTGCGCAAGCTTGGCTGCGGAGCGGCTGCTGATGCT 9470
QY 4453 CTTTGTCCAGATCATCTGATCGAAGACCGGCTTCCATCCGAGTACGCTGCTCGCTCGA 4512
DB 9469 CTTTGTCCAGATCATCTGATCGAAGACCGGCTTCCATCCGAGTACGCTGCTCGCTCGA 9410
QY 4513 TGGATGTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4572
DB 9409 TGGATGTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9350
QY 4573 GCATTTGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4632
DB 9349 GCATTTGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9290
QY 4633 CTTGCCCCGCACTTCCGCCAATAGCAGCGGCTTCCGCTTCCAGTGAACAGCTGCA 4692

Db 9289 CCTGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTCACAACTCGA 9230
Qy 4893 GCACAGCTGCGCAAGGAACGCCCGCTCGTGGCCACGATAGCCGGCTGCGCTCGTCT 4752
Db 9229 GCACAGCTGCGCAAGGAACGCCCGCTCGTGGCCACGATAGCCGGCTGCGCTCGTCT 9170
Qy 4753 GCAGTTTATTGAGGGCACCGGACACAGTTCGTTGACAAAAGAACCGGGCGCCCTGGG 4812
Db 9169 GCAGTTTATTGAGGGCACCGGACACAGTTCGTTGACAAAAGAACCGGGCGCCCTGGG 9110
Qy 4813 CTGACAGCCGGAACACCGCGGCATCAGAGCAGCCGATTGCTGTTGTGCCAGTCATAGC 4872
Db 9109 CTGACAGCCGGAACACCGCGGCATCAGAGCAGCCGATTGCTGTTGTGCCAGTCATAGC 9050
Qy 4873 CGAATAGCTCTCCACCCAGCCGCGGAGAACCTGGTGCAATCCATCTTGTTCATCA 4932
Db 9049 CGAATAGCTCTCCACCCAGCCGCGGAGAACCTGGTGCAATCCATCTTGTTCATCA 8990
Qy 4933 TGGGAAACGATCCTCATCCTGCTCTTGTGATCAGATCTTG 4971
Db 8989 TGGGAAACGATCCTCATCCTGCTCTTGTGATCAGATCTTG 8951

Search completed: February 15, 2004, 05:49:54
Job time : 350 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 05:29:07 ; Search time 1682 Seconds
(without alignments)
11569.936 Million cell updates/sec

Title: US-09-921-143-36
Perfect score: 5283
Sequence: 1 agcttgacctatgcgact.....tgatgcttgaggagcgtg 5283

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09E_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5283	100.0	5283	13	US-09-921-143-36
2	2240	42.4	8349	13	US-10-198-478-16
C 3	1793.4	33.9	4800	13	US-10-322-360-1
C 4	1692.4	32.0	5225	13	US-10-286-186-16
C 5	1632.8	30.9	5285	14	US-10-038-001-1
6	1603.8	30.4	2192	15	US-10-021-403A-10
7	1603.8	30.4	3534	12	US-10-395-709-11
8	1603.8	30.4	3534	12	US-10-395-709-12
9	1603.8	30.4	3534	12	US-10-395-709-13
10	1603.8	30.4	3534	12	US-10-395-709-14
11	1603.8	30.4	3534	12	US-10-395-709-15
12	1603.8	30.4	3534	15	US-10-021-403A-9
13	1600	30.3	5966	13	US-10-136-837-3
C 14	1571.6	29.7	4665	9	US-09-759-960-7
C 15	1570	29.7	4886	15	US-10-128-853-4

C 16	1570	29.7	5031	15	US-10-033-717-9	Sequence 9, Appl
C 17	1570	29.7	5054	15	US-10-033-717-8	Sequence 8, Appl
C 18	1570	29.7	5130	15	US-10-033-717-7	Sequence 7, Appl
C 19	1570	29.7	5168	15	US-10-033-717-5	Sequence 5, Appl
C 20	1569.8	29.7	4692	15	US-10-161-403-29	Sequence 29, Appl
C 21	1569.8	29.7	4727	13	US-10-115-987B-12	Sequence 12, Appl
C 22	1569.8	29.7	4733	11	US-09-797-496B-1	Sequence 1, Appl
C 23	1569.8	29.7	4733	13	US-10-177-390-1	Sequence 1, Appl
C 24	1569.8	29.7	4740	13	US-10-115-987B-13	Sequence 13, Appl
C 25	1569.8	29.7	4748	13	US-09-796-575-4	Sequence 4, Appl
C 26	1569.8	29.7	4862	15	US-10-161-403-87	Sequence 87, Appl
C 27	1569.8	29.7	4894	15	US-10-055-794-2	Sequence 2, Appl
C 28	1569.8	29.7	4992	13	US-09-796-575-5	Sequence 5, Appl
C 29	1569.8	29.7	5166	15	US-10-033-717-6	Sequence 6, Appl
C 30	1569.8	29.7	5953	12	US-10-157-317-13	Sequence 13, Appl
C 31	1569.8	29.7	5953	12	US-10-157-339-13	Sequence 13, Appl
C 32	1569.8	29.7	5953	13	US-10-157-305A-13	Sequence 13, Appl
C 33	1569.8	29.7	5953	13	US-10-157-391-13	Sequence 13, Appl
C 34	1569.8	29.7	5953	13	US-10-157-095-13	Sequence 13, Appl
C 35	1569.8	29.7	5953	13	US-10-157-219A-13	Sequence 13, Appl
C 36	1569.8	29.7	5953	13	US-10-157-309-13	Sequence 13, Appl
C 37	1569.8	29.7	5953	13	US-10-154-951B-13	Sequence 13, Appl
C 38	1569.8	29.7	5953	13	US-10-157-299-13	Sequence 13, Appl
C 39	1569.8	29.7	5953	13	US-10-156-831-13	Sequence 13, Appl
C 40	1569.8	29.7	5953	13	US-10-156-902-13	Sequence 13, Appl
C 41	1569.8	29.7	5953	13	US-10-157-147-13	Sequence 13, Appl
C 42	1569.8	29.7	5953	13	US-10-157-166-13	Sequence 13, Appl
C 43	1569.8	29.7	5953	13	US-10-157-318-13	Sequence 13, Appl
C 44	1569.8	29.7	5953	13	US-10-156-811-13	Sequence 13, Appl
C 45	1569.8	29.7	5953	13	US-10-157-320A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-921-143-36
; Sequence 36, Application US/09921143
; Publication No. US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P6
; CURRENT APPLICATION NUMBER: US/09/921,143
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 36
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-36

Query Match 100.0%; Score 5283; DB 13; Length 5283;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ARCTTGACCTTATGGACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTAT	60
Db	1	ARCTTGACCTTATGGACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTAT	60
Qy	61	TACCATGGTGATCGGTTTTGGCAGTACATCAATGGCGGTGGATCGGTTTGACTCACG	120
Db	61	TACCATGGTGATCGGTTTTGGCAGTACATCAATGGCGGTGGATCGGTTTGACTCACG	120
Qy	121	GGATTTCMAAGTCTCCACCCACTGAGTCAATGGAGTTTGTGTCACCAAAATCA	180
Db	121	GGATTTCMAAGTCTCCACCCACTGAGTCAATGGAGTTTGTGTCACCAAAATCA	180
Qy	181	ACGAGACTTTCACCAAAATGTCGTAAACATCCGCCCCCATTTGACCAATGGCGGTGAGGC	240

[illegible]

Db	3481	GGTAACAGGATTACACAGCGAGGATATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGG	3541
Qy	3541	CCTAACTACGGCTACACTAGAAACAAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGATT	3600
Db	3541	CCTAACTACGGCTACACTAGAAACAAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGATT	3600
Qy	3601	ACCTTCGGMAAAAGAGTTGGTAGCTCTTGATTCGGGMAAACAACACCGCTGGTAGCGGT	3660
Db	3601	ACCTTCGGMAAAAGAGTTGGTAGCTCTTGATTCGGGMAAACAACACCGCTGGTAGCGGT	3660
Qy	3661	GGTTTTTTTGTTCGAACAGCAGATATACCGCGCAGAAAAAAAAGATCTCAAGAAGATCCT	3720
Db	3661	GGTTTTTTTGTTCGAACAGCAGATATACCGCGCAGAAAAAAAAGATCTCAAGAAGATCCT	3720
Qy	3721	TTGATCTTTTCTACGGGCTCTGACGCTCAGTGAAACGAAACTCACGTTAAGGATTTTG	3780
Db	3721	TTGATCTTTTCTACGGGCTCTGACGCTCAGTGAAACGAAACTCACGTTAAGGATTTTG	3780
Qy	3781	GTCAATGAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCACTCCCTGCACTGCTCGGG	3840
Db	3781	GTCAATGAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCACTCCCTGCACTGCTCGGG	3840
Qy	3841	GCATGGATCGCGGATACCGCTGCTGGTTTCTGGATGCCAGGATTTGCATGCCCG	3900
Db	3841	GCATGGATCGCGGATACCGCTGCTGGTTTCTGGATGCCAGGATTTGCATGCCCG	3900
Qy	3901	TAGAACTCCCGGAGGTGTCACGCTCAGCAGCAGCTGAAACCACTCGGAGGGGATCG	3960
Db	3901	TAGAACTCCCGGAGGTGTCACGCTCAGCAGCAGCTGAAACCACTCGGAGGGGATCG	3960
Qy	3961	AGCCCGGGTGGGCGAAGAACTCCAGCATCAGATCCCCGCTGGAGGATCATCCAGCCG	4020
Db	3961	AGCCCGGGTGGGCGAAGAACTCCAGCATCAGATCCCCGCTGGAGGATCATCCAGCCG	4020
Qy	4021	GGCTCCCGGAAAACGATTCGGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAA	4080
Db	4021	GGCTCCCGGAAAACGATTCGGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAA	4080
Qy	4081	TCTCGTATGCGAGTTGGGCTCGCTGGTGGTTCATTTGAAACCCAGAGTCCCGCTC	4140
Db	4081	TCTCGTATGCGAGTTGGGCTCGCTGGTGGTTCATTTGAAACCCAGAGTCCCGCTC	4140
Qy	4141	AGAAGAACTCGTCAAGAAGSCGATAGAAGCGATGCGCTGCGAAATCGGAGCGCGATAC	4200
Db	4141	AGAAGAACTCGTCAAGAAGSCGATAGAAGCGATGCGCTGCGAAATCGGAGCGCGATAC	4200
Qy	4201	CGTAAAGCAGAGAGCGGTACGCCATTCGGCCGCAAGCTCTTCAGCAATATCACGGG	4260
Db	4201	CGTAAAGCAGAGAGCGGTACGCCATTCGGCCGCAAGCTCTTCAGCAATATCACGGG	4260
Qy	4261	TAGCCAAAGCTATGTCTCTGATAGCGGTCCGCCACACCCAGCGGCACAGTCGATGAATC	4320
Db	4261	TAGCCAAAGCTATGTCTCTGATAGCGGTCCGCCACACCCAGCGGCACAGTCGATGAATC	4320
Qy	4321	CAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTACGA	4380
Db	4321	CAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTACGA	4380
Qy	4381	CGAGATCTTCGGCTCGGGCATCGCGCTTCGAGCTGCGCAACAGTTCGGCTGGCGGA	4440
Db	4381	CGAGATCTTCGGCTCGGGCATCGCGCTTCGAGCTGCGCAACAGTTCGGCTGGCGGA	4440
Qy	4441	GCCCTCGATGCTCTTCGTCCAGATCATCTGATCGAACAGACCGGCTTCATCCGAGTAC	4500
Db	4441	GCCCTCGATGCTCTTCGTCCAGATCATCTGATCGAACAGACCGGCTTCATCCGAGTAC	4500
Qy	4501	GTGCTCGCTCGATGCCATGTTTCGTTTGTGTCGAATGGGCAAGTAGCCGATCACGG	4560
Db	4501	GTGCTCGCTCGATGCCATGTTTCGTTTGTGTCGAATGGGCAAGTAGCCGATCACGG	4560
Qy	4561	TATGCAAGCGCCGATTCGATCAGCCATGATGGATCTTTCTCGGCAAGGAGCAAGGTGAG	4620

Db 4561 TATGACGCCGCCGCAATTGCATACGCCGATGATGATCTTTTCGGCAGGAGCAAGTGAG 4620
Qy 4621 ATGACAGGAGATCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAG 4680
Db 4621 ATGACAGGAGATCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAG 4680
Qy 4681 TGACAACTGAGCAGCAGCTGCGCAAGAGAACGCCCTGCTGGCAGCCACGATAGCCGG 4740
Db 4681 TGACAACTGAGCAGCAGCTGCGCAAGAGAACGCCCTGCTGGCAGCCACGATAGCCGG 4740
Qy 4741 CTGCTCTGCTGCTGAGTTCAATTCAGGCGCACCGCACAGGTGGTCTTGACAAAAGAACCG 4800
Db 4741 CTGCTCTGCTGCTGAGTTCAATTCAGGCGCACCGCACAGGTGGTCTTGACAAAAGAACCG 4800
Qy 4801 GGGCCCCCTGCTGACAGCCGGAACACGCGCGGCATCAGAGCAGCCGATGTCTGTGTG 4860
Db 4801 GGGCCCCCTGCTGACAGCCGGAACACGCGCGGCATCAGAGCAGCCGATGTCTGTGTG 4860
Qy 4861 CCCAGTCATAGCCGAATAGCTCTCCACCCAGCGCGGAGAACCTGGGTGCAATCCAT 4920
Db 4861 CCCAGTCATAGCCGAATAGCTCTCCACCCAGCGCGGAGAACCTGGGTGCAATCCAT 4920
Qy 4921 CTTGTTCAATCATGCGAAACGATCCTCATCTCTGTCTCTTGATCAGATCTTGTCCCTGC 4980
Db 4921 CTTGTTCAATCATGCGAAACGATCCTCATCTCTGTCTCTTGATCAGATCTTGTCCCTGC 4980
Qy 4981 GCCATCAGATCTTGGCGCAAGAACCCATCAGTTTACTTTGAGGSCCTTCCCACT 5040
Db 4981 GCCATCAGATCTTGGCGCAAGAACCCATCAGTTTACTTTGAGGSCCTTCCCACT 5040
Qy 5041 TACAGAGGCGGCCCGCAGCTGGCAATTCGGGTCGCTTGCTGCCATATAAACCGGCCAGT 5100
Db 5041 TACAGAGGCGGCCCGCAGCTGGCAATTCGGGTCGCTTGCTGCCATATAAACCGGCCAGT 5100
Qy 5101 CTAGCTATCCCATGTAAGCCCATGCAAGCTACTGCTTTCTTTTGGCTTGGCTTTT 5160
Db 5101 CTAGCTATCCCATGTAAGCCCATGCAAGCTACTGCTTTCTTTTGGCTTGGCTTTT 5160
Qy 5161 CCCTTGTCAGATAGCCCGCAGTGTGACATTCATCCGGGTGACGACCGTTTCTGGGAC 5220
Db 5161 CCCTTGTCAGATAGCCCGCAGTGTGACATTCATCCGGGTGACGACCGTTTCTGGGAC 5220
Qy 5221 TGGCTTTCTAGCTTCCGCTTCTTTAGCAGCCCTTGGCCCTGAGTCTTGGCGAGC 5280
Db 5221 TGGCTTTCTAGCTTCCGCTTCTTTAGCAGCCCTTGGCCCTGAGTCTTGGCGAGC 5280
Qy 5281 GTG 5283
Db 5281 GTG 5283

RESULT 2
US-10-198-478-16
; Sequence 16, Application US/10198478
; Publication No. US20030189336A1
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; TITLE OF INVENTION: Improved Methods for Transforming Plants to Express delta-Endotoxin
; FILE REFERENCE: 38-21 (13547) B
; CURRENT APPLICATION NUMBER: US/10/198,478
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/186, 002
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 8349
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: completely synthesized
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (1) (8349)
; OTHER INFORMATION:
US-10-198-478-16

Query Match 42.4%; Score 2240; DB 13; Length 8349;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 25; Indels 144; Gaps 1;

Qy 2706 CTAGACGCTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCC 2765
Db 5911 CTTGGCGTATCATGGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCC 5970
Qy 2766 ACACAAATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGATGAGCTA 2825
Db 5971 ACACAAATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGATGAGCTA 6030
Qy 2826 ACTCACATTAATTCGCTTGGCTCACTCGCCGCTTTCCAGTCGGAACCTGCTGGTCCA 2885
Db 6031 ACTCACATTAATTCGCTTGGCTCACTCGCCGCTTTCCAGTCGGAACCTGCTGGTCCA 6090
Qy 2886 GCTGCATTAATGAATTCGCGCAACGCGCGGAGAGCGGTTTTCGTAATTCGGGCTCTTC 2945
Db 6091 GCTGCATTAATGAATTCGCGCAACGCGCGGAGAGCGGTTTTCGTAATTCGGGCTCTTC 6150
Qy 2946 CGCTTCTCTGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
Db 6151 CGCTTCTCTGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6210
Qy 3006 TCACTCAAGGCGGTAATACGCTTTATCCACAGATCAGGGGTAACACGAGGAAAGACAT 3065
Db 6211 TCACTCAAGGCGGTAATACGCTTTATCCACAGATCAGGGGTAACACGAGGAAAGACAT 6270
Qy 3066 GTGAGCAAGAGGCGAGCAAAAGGCGAGGAAACCGTAAAGAGCGCGTTCGCTGGGTTTTT 3125
Db 6271 GTGAGCAAGAGGCGAGCAAAAGGCGAGGAAACCGTAAAGAGCGCGTTCGCTGGGTTTTT 6330
Qy 3126 CCAATAGCTCCGCCCCCTGACGAGCATCAAAATATCGACGCTCAAGTCAGAGGTGGCG 3185
Db 6331 CCAATAGCTCCGCCCCCTGACGAGCATCAAAATATCGACGCTCAAGTCAGAGGTGGCG 6390
Qy 3186 AAACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTTGGAGGCTCCCTCGTGGGCTC 3245
Db 6391 AAACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTTGGAGGCTCCCTCGTGGGCTC 6450
Qy 3246 TCCGTGTTCCGACCCGCTGACCGGATACCTGCGCGCTTCTCCCTTCCGCTGGGAGCGT 3305
Db 6451 TCCGTGTTCCGACCCGCTGACCGGATACCTGCGCGCTTCTCCCTTCCGCTGGGAGCGT 6510
Qy 3306 GCGCTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTGTAGGTGCTTCGCTCAA 3365
Db 6511 GCGCTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTGTAGGTGCTTCGCTCAA 6570
Qy 3366 GCTGGCTGTGTGACGAACCCCGCTTACGCCGAGCGCTGGCTTATCCGCTAACTA 3425
Db 6571 GCTGGCTGTGTGACGAACCCCGCTTACGCCGAGCGCTGGCTTATCCGCTAACTA 6630
Qy 3426 TCGCTTTGAGTCCAAACCCCGGTAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA 3485
Db 6631 TCGCTTTGAGTCCAAACCCCGGTAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA 6690
Qy 3486 CAGGATTAGCAGAGGAGGTATGTAGCGGTGTGTACAGAGTTCTTGAAGTGGTGGCTTAA 3545
Db 6691 CAGGATTAGCAGAGGAGGTATGTAGCGGTGTGTACAGAGTTCTTGAAGTGGTGGCTTAA 6750
Qy 3546 CTACGGCTTACTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGGTACTT 3605
Db 6751 CTACGGCTTACTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGGTACTT 6810
Qy 3606 CGGAAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTT 3665
Db 6811 CGGAAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTT 6870

RESULT 7

```

US-10-395-709-11
; Sequence 11, Application US/10395709
; Publication NO. US20040014645A1
; GENERAL INFORMATION:
; APPLICANT: Advivys
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE POL
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE POL
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE POL
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE POL
; FILE REFERENCE: 108328.00115 - AVSI-0021P1
; CURRENT APPLICATION NUMBER: US/10/395.709
; CURRENT FILING DATE: 2003-03-24
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence for the operatively linked components of the HV-GHRH pla
; OTHER INFORMATION: smid.
US-10-395-709-11

```

Query Match 30.4%; Score 1603.8; DB 12; Length 3534;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;

Qy	2706	CTAGACGTAATCATGGTTCATAGCTGTTTCTGTGTGAAATGTATTCCGCTCAATTC	2765
Db	1412	CTTGGCGTAATCATGGTTCATAGCTGTTTCTGTGTGAAATGTATTCCGCTCAATTC	1471
Qy	2766	ACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTCCCTAAATGAGTGAAGCTA	2825
Db	1472	ACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTCCCTAAATGAGTGAAGCTA	1531
Qy	2826	ACTCACATTAATTGCGTTTGCGCTCACTGCCGCGCTTTCAGTTCGGGAAACCTCTGCTGCCA	2885
Db	1532	ACTCACATTAATTGCGTTTGCGCTCACTGCCGCGCTTTCAGTTCGGGAAACCTCTGCTGCCA	1591
Qy	2886	GCTGCATTAAATCGGCCTAAACGCGCGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTC	2945
Db	1592	GCTGCATTAAATCGGCCTAAACGCGCGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTC	1651
Qy	2946	CGCTTCCTCGTCACTGACTCGCTCGCTCGTTCGGTTCGGGTCGGCGAGCGGTATCAGC	3005
Db	1652	CGCTTCCTCGTCACTGACTCGCTCGCTCGTTCGGTTCGGGTCGGCGAGCGGTATCAGC	1711
Qy	3006	TCACCTCAAGCGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGACGAAAGACAT	3065
Db	1712	TCACCTCAAGCGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGACGAAAGACAT	1771
Qy	3066	GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTCGTGCGGTTTTT	3125
Db	1772	GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTCGTGCGGTTTTT	1831
Qy	3126	CCATAGGCTCCGCCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGGC	3185
Db	1832	CCATAGGCTCCGCCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGGC	1891
Qy	3186	AAACCCGACAGGACTATAAAGATACACAGCGCTTTCCCGCTGGAAAGCTCCCTCTGTCGCGTC	3245
Db	1892	AAACCCGACAGGACTATAAAGATACACAGCGCTTTCCCGCTGGAAAGCTCCCTCTGTCGCGTC	1951
Qy	3246	TCCTGTTCCGACCTTCCGGCTTACGGATACCTGTGCGCTTTCTCCGTTCCGGGAAGCGT	3305
Db	1952	TCCTGTTCCGACCTTCCGGCTTACGGATACCTGTGCGCTTTCTCCGTTCCGGGAAGCGT	2011
Qy	3306	GGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGCTCAA	3365
Db	2012	GGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGCTCAA	2071
Qy	3366	GCTGGGCTGTGTGACGAACCCCGGTTTCAGCCCGACCGCTGCGGCTTATCCGGTAACATA	3425
Db	2072	GCTGGGCTGTGTGACGAACCCCGGTTTCAGCCCGACCGCTGCGGCTTATCCGGTAACATA	2131

QY 3846 GATCGCGGATAGCCGCTGCTGCTTTCTTGATGCGGACGGATTTGCACTGCGGTAGAA 3905
Db 2450 ----- 2449
QY 3906 CTCCGCGAGGTGCTCAGCCTCAGGCAGCAGCTGACCACTCGCGGAGGGGATCGAGCCC 3965
Db 2450 ----- 2449
QY 3966 GGGGTGGGGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGCTC 4025
Db 2450 ----- 2449
QY 4026 CCGGAAAACGATTCGGAAGCCCACTTTTCATAGAAGCGCGGTGGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
QY 4086 TGATGCGAGGTGGCGCTGCTTGCTGCTGCTATTTGCAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 -----CGCTCAGAAG 2459
QY 4146 AACTCTGAAGAGCGATAGAGCGGATGCGCTGCGAATCGGAGCGGCGATACCGTAA 4205
Db 2460 AACTCTGAAGAGCGATAGAGCGGATGCGCTGCGAATCGGAGCGGCGATACCGTAA 2519
QY 4206 AGCAGGAGGAGCGGTACGCCATTCGCCGCAAGCTCTTACGCAATATCACGGGTAGCC 4265
Db 2520 AGCAGGAGGAGCGGTACGCCATTCGCCGCAAGCTCTTACGCAATATCACGGGTAGCC 2579
QY 4266 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGATCAAGAA 4325
Db 2580 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGATCAAGAA 2639
QY 4326 AAGCGGCCATTTCCACCATGATATTCGCAAGCAGGATCGCATGGTCAACGAGAGA 4385
Db 2640 AAGCGGCCATTTCCACCATGATATTCGCAAGCAGGATCGCATGGTCAACGAGAGA 2699
QY 4386 TCCTCGCGCTCGGCGATCGCGCTTGAGCCTGCGGAAACAGTTTCGCTGCGCGGAGCGCC 4445
Db 2700 TCCTCGCGCTCGGCGATCGCGCTTGAGCCTGCGGAAACAGTTTCGCTGCGCGGAGCGCC 2759
QY 4446 TGATGCTCTTGCTCAGATATCTCTGATCGACAGACCGGCTTCATTCGAGTACGTGCT 4505
Db 2760 TGATGCTCTTGCTCAGATATCTCTGATCGACAGACCGGCTTCATTCGAGTACGTGCT 2819
QY 4506 CGCTCGATGCGATGTTTCGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4565
Db 2820 CGCTCGATGCGATGTTTCGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2879
QY 4566 AGCGCGCGCATTCGATCAGCCTATGATGATGATGATGATGATGATGATGATGATGATGAT 4625
Db 2880 AGCGCGCGCATTCGATCAGCCTATGATGATGATGATGATGATGATGATGATGATGATGAT 2939
QY 4626 AGGAGATCTGCGCGGCACTTCGCCCAATAGCAGCAGCTGCTTCCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTGCGCGGCACTTCGCCCAATAGCAGCAGCTGCTTCCCGCTTCAGTGACA 2999
QY 4686 ACCTCGAGCAGCTGCGCAAGGAACCGCGTCTGTCGCGCAGGACGATAGCGCGCTGCC 4745
Db 3000 ACCTCGAGCAGCTGCGCAAGGAACCGCGTCTGTCGCGCAGGACGATAGCGCGCTGCC 3059
QY 4746 TCGTCTGCGATTCATTCAGGCGACCGGACAGCTGCTTTCGCAAAAAGAACCGGGGCG 4805
Db 3060 TCGTCTGCGATTCATTCAGGCGACCGGACAGCTGCTTTCGCAAAAAGAACCGGGGCG 3119
QY 4806 CCCTGCGCTCAGACCGGCAACCGCGGATCAGACGACCGATTCGCTGTTGTCGCCAG 4865
Db 3120 CCCTGCGCTCAGACCGGCAACCGCGGATCAGACGACCGATTCGCTGTTGTCGCCAG 3179
QY 4866 TCATAGCCGAATAGCCTTCCACCCAGCGCGCGAGAACCTTCGCTGCGAATCCATCTTGT 4925
Db 3180 TCATAGCCGAATAGCCTTCCACCCAGCGCGCGAGAACCTTCGCTGCGAATCCATCTTGT 3239

QY 4926 TCAATCATGCGAAACGATCCTCATCTCTGCTCTTGTATCATGATCTTGTATCCCTGCGCAT 4985
Db 3240 TCAATCATGCGAAACGATCCTCATCTCTGCTCTTGTATCATGATCTTGTATCCCTGCGCAT 3299
QY 4986 CAGATCTCTGCGGCAAGAACGATCCAGTTTACTTTTTCAGGCTTCCCAACCTTACCA 5045
Db 3300 CAGATCTCTGCGGCAAGAACGATCCAGTTTACTTTTTCAGGCTTCCCAACCTTACCA 3359
QY 5046 GAGGCGCCCGCAGCTGCGAATTCGGTTCGCTTGTGTCATATAAACCGCCAGTCTAGC 5105
Db 3360 GAGGCGCCCGCAGCTGCGAATTCGGTTCGCTTGTGTCATATAAACCGCCAGTCTAGC 3419
QY 5106 TATCGCCATGTAAGCCACTGCAAGTACCTGCTTCTCTT 5146
Db 3420 AACTGTTGGAAGGCGATCGGTGCGGCGCTCTTCGCTATT 3460

RESULT 9
US-10-395-709-13
; Sequence 13, Application US/10395709
; Publication No. US20040014645A1
; GENERAL INFORMATION:
; APPLICANT: ADVISYS
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE F
; FILE REFERENCE: GLUTAMATE ("PLG") SYSTEM
; CURRENT APPLICATION NUMBER: US/10395,709
; CURRENT FILING DATE: 2003-03-24
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence for the operatively linked components of the TV-GHRH F
; OTHER INFORMATION: smid.
US-10-395-709-13

Query Match 30.4%; Score 1603.8; DB 12; Length 3534;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;
QY 2706 CTAGACGTAATCATGCTATGCTATGCTTCTGCTGTAATTTATTCGCTCAATTC 2765
Db 1412 CTGCGCTAATCATGCTATGCTATGCTTCTGCTGTAATTTATTCGCTCAATTC 1471
QY 2766 ACACAAATACGACCGCGAAGCATTAAGTGAAGCTGCGGTGCTTAATGAGTACGTA 2825
Db 1472 ACACAAATACGACCGCGAAGCATTAAGTGAAGCTGCGGTGCTTAATGAGTACGTA 1531
QY 2826 ACTCAGATTAATGCTGCTGCTCACTGCGCTTTCAGTCCGGAACCTGCTGCGCA 2885
Db 1532 ACTCAGATTAATGCTGCTGCTCACTGCGCTTTCAGTCCGGAACCTGCTGCGCA 1591
QY 2886 GCTGCTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2945
Db 1592 GCTGCTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1651
QY 2946 GCTTTCCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3005
Db 1652 GCTTTCCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
QY 3006 TCATCAAAAGCGGTATATACGGTTATCCACAGAAATCAGGGGATACCGCAGAAAGAACAT 3065
Db 1712 TCATCAAAAGCGGTATATACGGTTATCCACAGAAATCAGGGGATACCGCAGAAAGAACAT 1771
QY 3066 GTGAGCAAAAGCGGACAGCAAAAGCGGACCGGTAAAGCGCGCTTGTGCGGTGTTT 3125
Db 1772 GTGAGCAAAAGCGGACAGCAAAAGCGGACCGGTAAAGCGCGCTTGTGCGGTGTTT 1831
QY 3126 CCATAGGCTCGCGCCCTTCGACGAGCATCAAAAATTCAGCTCAAGTCAGAGTGGCG 3185

Db 1832 CCATAGGCTCGCCCCCTGACGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGGCG 1891
Qy 3186 AAACCCGACAGGACTATAAGATACACAGGGTTCCTCCCTCGGAAGTCTCCCTCGTGGCTC 3245
Db 1892 AAACCCGACAGGACTATAAGATACACAGGGTTCCTCCCTCGGAAGTCTCCCTCGTGGCTC 1951
Qy 3246 TCCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGT 3305
Db 1952 TCCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGT 2011
Qy 3306 GCGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAA 3365
Db 2012 GCGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAA 2071
Qy 3366 GCTGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAAC 3425
Db 2072 GCTGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAAC 2131
Qy 3426 TCCTCTTGTAGTCCAAACCGGTAAACACGACTTATCGCCACTGGCAGCCACTGTTAA 3485
Db 2132 TCGTCTTGTAGTCCAAACCGGTAAACACGACTTATCGCCACTGGCAGCCACTGTTAA 2191
Qy 3486 CAGATTAGCAGAGCGAGGTATGAGCGGTGTCTACAGAGTCTTGAAGTGTGGCTAA 3545
Db 2192 CAGATTAGCAGAGCGAGGTATGAGCGGTGTCTACAGAGTCTTGAAGTGTGGCTAA 2251
Qy 3546 CTAGGCTTACACTAGAGACAGTATTTGTTGTTCTGCTCTGCTGAGCCAGTACCTT 3605
Db 2252 CTAGGCTTACACTAGAGACAGTATTTGTTGTTCTGCTCTGCTGAGCCAGTACCTT 2311
Qy 3606 CGGAAAGAGTGTGTAGCTTCTGATCCGGCAAAACAAACACCGCTGTTAGCGGTGTTT 3665
Db 2312 CGGAAAGAGTGTGTAGCTTCTGATCCGGCAAAACAAACACCGCTGTTAGCGGTGTTT 2371
Qy 3666 TTTTGTGTAAGCAGCAGATTAGCGCAGAAAGAGATCTCAAGAGATCTTTGAT 3725
Db 2372 TTTTGTGTAAGCAGCAGATTAGCGCAGAAAGAGATCTCAAGAGATCTTTGAT 2431
Qy 3726 CTTTCTACGGGCTGACGCTCAGTGGAGCAAAACTCAGTTAAGGATTTTGTTCAT 3785
Db 2432 CTTTCTACGGGCTGACGCTCAGTGGAGCAAAACTCAGTTAAGGATTTTGTTCAT 3845
Qy 3786 GAGATTATGTCGACCAAGCGGCGCATCGTGCCTCCCACTCTCTGCAAGTTTCGGGGGATG 2449
Db 2450 ----- 2449
Qy 3846 GATCGCGGATAGCGGCTGCTGTTTCTGATGCGGAGGATTTGCACTGCGGCTAGAA 3905
Db 2450 ----- 2449
Qy 3906 CTCGCGAGGTGTCAGCTTACGCGACAGCTGAACCACTCGCGAGGGATCGAGCC 3965
Db 2450 ----- 2449
Qy 3966 GGGGTGGGGAAGAACTCCAGATGAGTCCCGCGCTGGAGGATCATCCAGCGGGCTC 4025
Db 2450 ----- 2449
Qy 4026 CCGGAAACGATTCGGAAGCCCAACCTTTTCATAGAGCGCGGTGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
Qy 4086 TGATGGCAGGTTGGCGCTCGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 -----CGTCTCAAG 2459
Qy 4146 AACTCGTCAGAAGCGATAGAAGCGATGGCTGCGAATCGGAGCGGCGATACCGTAA 4205
Db 2460 AACTCGTCAGAAGCGATAGAAGCGATGGCTGCGAATCGGAGCGGCGATACCGTAA 2519
Qy 4206 AGCAGAGAGCGGTGAGCCATTCGCGCAAGCTCTTCAGCAATATCAGGGGTAGCC 4265
Db 2520 AGCAGAGAGCGGTGAGCCATTCGCGCAAGCTCTTCAGCAATATCAGGGGTAGCC 2579

Qy 4266 AACGCTATGCTCTGATAGCGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAA 4325
Db 2580 AACGCTATGCTCTGATAGCGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAA 2639
Qy 4326 AAGCGCCATTTTCCACATGATATTCGGCAAGCAGGATCCCATGGTTCACGACGAGA 4385
Db 2640 AAGCGCCATTTTCCACATGATATTCGGCAAGCAGGATCCCATGGTTCACGACGAGA 2699
Qy 4386 TCCTCGCGCTCGGCGATGCGCGCTTGAAGCTGCGCAACAGTTCGGCTGGCGAGGCC 4445
Db 2700 TCCTCGCGCTCGGCGATGCGCGCTTGAAGCTGCGCAACAGTTCGGCTGGCGAGGCC 2759
Qy 4446 TGATGCTCTTCTGTCAGATCATCTGATCGACAGACCGGCTTCATCCGATGAGTGC 4505
Db 2760 TGATGCTCTTCTGTCAGATCATCTGATCGACAGACCGGCTTCATCCGATGAGTGC 2819
Qy 4506 CGCTCGATGCGATGTTTCGCTTGGTGGTTCGAAATGGCAGGTAGCGGATCAAGCGTATGC 4565
Db 2820 CGCTCGATGCGATGTTTCGCTTGGTGGTTCGAAATGGCAGGTAGCGGATCAAGCGTATGC 2879
Qy 4566 AGCGCGCCATTCGATCAGCATGATGATCTTCTCGGACGAGCAAGGTGAGTGC 4625
Db 2880 AGCGCGCCATTCGATCAGCATGATGATCTTCTCGGACGAGCAAGGTGAGTGC 2939
Qy 4626 AGGAGATCTTCCCGCGCACTTCCGCAATAGCAGCCAGTCCCTTCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTTCCCGCGCACTTCCGCAATAGCAGCCAGTCCCTTCCGCTTCAGTGACA 2999
Qy 4686 AGCTCGACACAGCTGCGCAAGAAACGCGCTGTCGCGCAGCCACGATAGCCGCGTGC 4745
Db 3000 AGCTCGACACAGCTGCGCAAGAAACGCGCTGTCGCGCAGCCACGATAGCCGCGTGC 3059
Qy 4746 TCGTCTCTCAGTTCATTCAGGCGCACCGGACAGTTCGTTGACAAAAAGAACCGGCGC 4805
Db 3060 TCGTCTCTCAGTTCATTCAGGCGCACCGGACAGTTCGTTGACAAAAAGAACCGGCGC 3119
Qy 4806 CCCTCGCTGACAGCGGAAACACGCGGCGATCAGACAGCGGATGTTGTGTGCCAG 4865
Db 3120 CCCTCGCTGACAGCGGAAACACGCGGCGATCAGACAGCGGATGTTGTGTGCCAG 3179
Qy 4866 TCATAGCCGATAGCTTCCACCCAGCGCGGAGAACCTGCTGCAATCCATCTGT 4925
Db 3180 TCATAGCCGATAGCTTCCACCCAGCGCGGAGAACCTGCTGCAATCCATCTGT 3239
Qy 4926 TCATATCATCGGAAACGATCCCTCATCTCTCTTGTATCAGATCTTGTCCCTCGCCAT 4985
Db 3240 TCATATCATCGGAAACGATCCCTCATCTCTCTTGTATCAGATCTTGTCCCTCGCCAT 3299
Qy 4986 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCGACCTTACCA 5045
Db 3300 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCGACCTTACCA 3359
Qy 5046 GAGGCGGCCAGCTGGCAATTCGGTTCGCTTCTGTCATATAAAGCCGCGCTAGC 5105
Db 3360 GAGGCGGCCAGCTGGCAATTCGGTTCGCTTCTGTCATATAAAGCCGCGCTAGC 3419
Qy 5106 TATGCCCATGTAAGCCCACTGCAAGTACCTGCTTCTCTT 5146
Db 3420 AACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCTGCTATT 3460

RESULT 10

US-10-395-709-14
; Sequence 14, Application US/10395709
; Publication No. US20040014645A1
; GENERAL INFORMATION:
; APPLICANT: Advisys
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE I
; TITLE OF INVENTION: GLUTAMATE ("PLG") SYSTEM
; FILE REFERENCE: 108328.00115 - AVSI-0021P1
; CURRENT APPLICATION NUMBER: US/10/395,709
; CURRENT FILING DATE: 2003-03-24

3306 GCGCTTTCTCAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGGTCTCCAA 3365
Db
2012 GCGGCTTTCTCAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGGTCTCCAA 2071
Qy
3366 GCTGGGCTGTGTCAGCAAGCCCGCTCAGCCCGACCGCTGCGCTTATCCGGTAACTA 3425
Db
2072 GCTGGGCTGTGTCAGCAAGCCCGCTCAGCCCGACCGCTGCGCTTATCCGGTAACTA 2131
Qy
3426 TCGTCTTGTAGTCAACCCCGTAAAGACAGCACTTATCGCCACTGCGAGCAGCCACTGGTAA 3485
Db
2132 TCGTCTTGTAGTCAACCCCGTAAAGACAGCACTTATCGCCACTGCGAGCAGCCACTGGTAA 2191
Qy
3486 CAGGATTTAGCAGAGCGAGGTATCTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAA 3545
Db
2192 CAGGATTTAGCAGAGCGAGGTATCTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAA 2251
Qy
3546 CTACGGCTACACTAGAGAAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 3605
Db
2252 CTACGGCTACACTAGAGAAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 2311
Qy
3606 CGGAAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTT 3665
Db
2312 CGGAAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTT 2371
Qy
3666 TTTTGTGTTGCAAGCAGCAGATTAAGCGCAGAAAAAAGGATCTCAAGAAAGATCCCTTGAT 3725
Db
2372 TTTTGTGTTGCAAGCAGCAGATTAAGCGCAGAAAAAAGGATCTCAAGAAAGATCCCTTGAT 2431
Qy
3726 CTTTCTACGGGCTGTCAGCCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGGTCAT 3785
Db
2432 CTTTCTACGGGCTGTCAGCCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGGTCAT 2449
Qy
3786 GAGATTATCGTCGACCAAGCGGCATCGTGGCTCCCACTCTCGAGTTGCGGGGCATG 3845
Db
2450 ----- 2449
Qy
3846 GATCGCGGATAGCGCTGCTGCTTCTCGATGCCGACGATTTCCACTGCGGGTAGAA 3905
Db
2450 ----- 2449
Qy
3906 CTCGCGAGGTGCTCAGCCTCAGGAGCAGCTGAACCAACTCGCGAGGGATCGAGCCC 3965
Db
2450 ----- 2449
Qy
3966 GGGGTGGCGAAGAACTCCAGCATGATCCCGCGCTGAGGATCATCCAGCGCGCTC 4025
Db
2450 ----- 2449
Qy
4026 CCGGAAAAAGATTCCGAAGCCCAACCTTTCTATGAGAGCGCGGTGGAATCTCG 4085
Db
2450 ----- 2449
Qy
4086 TGATGCGAGTTGGGCGTCTGCTGGTCAATTTGCAACCCAGAGTCCCGCTCAGAAG 4145
Db
2450 -----CGCTCAGAAG 2459
Qy
4146 AACTCGTCAGAAGCGGATAGAAGCGATGCGTGAATCGGAGCGGCGATACCGTAA 4205
Db
2460 AACTCGTCAGAAGCGGATAGAAGCGATGCGTGAATCGGAGCGGCGATACCGTAA 2519
Qy
4206 AGCAGGAGAGCGGTGAGCCATTCGCGCCAGCTCTTCAGCAATATCAGCGGTAGCC 4265
Db
2520 AGCAGGAGAGCGGTGAGCCATTCGCGCCAGCTCTTCAGCAATATCAGCGGTAGCC 2579
Qy
4266 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCGCACAGTGTGAATCCAGAA 4325
Db
2580 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCGCACAGTGTGAATCCAGAA 2639
Qy
4326 AAGCGGCCATTTCCACCATGATATTCGCGAGCAGCATCGCCATGGGTACAGCAGA 4385
Db
2640 AAGCGGCCATTTCCACCATGATATTCGCGAGCAGCATCGCCATGGGTACAGCAGA 2699
Qy
4386 TCCTCGCGGTGGGCGATCGCGCTTGGAGCCTGGCGAAACAGTTTCGGTGGCGAGCCCC 4445

2700 TCCTCGCGCTGGGCGATGCGCGCTTGGCTGGCAACAGTTTCGGTGGCGAGCCCC 2759
Qy
4446 TGATGCTCTTTCGTCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCT 4505
Db
2760 TGATGCTCTTTCGTCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCT 2819
Qy
4506 CGCTCGATGCGATGTTTCGCTTGGTTCGAATGGGCGAGTACCGGATCAAGGATATGC 4565
Db
2820 CGCTCGATGCGATGTTTCGCTTGGTTCGAATGGGCGAGTACCGGATCAAGGATATGC 2879
Qy
4566 AGCCCGCGCATTTGATCAGCCATGATGATACCTTTCGCGCAGGAGCAAGGTGAGATGAC 4625
Db
2880 AGCCCGCGCATTTGATCAGCCATGATGATACCTTTCGCGCAGGAGCAAGGTGAGATGAC 2939
Qy
4626 AGGAGATCTTCCCGCGCATTTGCGCCANTAGACCGAGTCCCTTCCGCTTCAGTGACA 4685
Db
2940 AGGAGATCTTCCCGCGCATTTGCGCCANTAGACCGAGTCCCTTCCGCTTCAGTGACA 2999
Qy
4686 ACCTCGAGCAAGCTGCGCAAGGAACCGCGTGTGGCCAGCCACGATAGCCGCGCTGCC 4745
Db
3000 ACCTCGAGCAAGCTGCGCAAGGAACCGCGTGTGGCCAGCCACGATAGCCGCGCTGCC 3059
Qy
4746 TCGTCTTGCATTCATTTCAAGGCAACCGGACAGTTCGCTTTGACAAAAAGAACCGGCGC 4805
Db
3060 TCGTCTTGCATTCATTTCAAGGCAACCGGACAGTTCGCTTTGACAAAAAGAACCGGCGC 3119
Qy
4806 CCCTGCGCTGACGCGGACACGCGGCGCATCGAGCAGCGGATTCGTTGTGCCAG 4865
Db
3120 CCCTGCGCTGACGCGGACACGCGGCGCATCGAGCAGCGGATTCGTTGTGCCAG 3179
Qy
4866 TCATAGCGGATAGCTCTCCACCCAGCGCGGAGAACCTGCGTCAATCCATCTTGT 4925
Db
3180 TCATAGCGGATAGCTCTCCACCCAGCGCGGAGAACCTGCGTCAATCCATCTTGT 3239
Qy
4926 TCATATGCGGAACGATCTCTATCTCTGTGATCAGATCTTGTATCCCTCGCGCAT 4985
Db
3240 TCATATGCGGAACGATCTCTATCTCTGTGATCAGATCTTGTATCCCTCGCGCAT 3299
Qy
4986 CAGATCTTGGCGGCAAGAACCCATCCAGTTTACTTTGAGGGCTTCCCAACCTTACCA 5045
Db
3300 CAGATCTTGGCGGCAAGAACCCATCCAGTTTACTTTGAGGGCTTCCCAACCTTACCA 3359
Qy
5046 GAGGCGCGCCAGCTGCGAATTCGGTTCGCTGCTGTCTATATAAACCGCGCGCTAGC 5105
Db
3360 GAGGCGCGCCAGCTGCGAATTCGGTTCGCTGCTGTCTATATAAACCGCGCGCTAGC 3419
Qy
5106 TATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTT 5146
Db
3420 AACTTTGGGAGGGCGATCGTGGCGCTCTTCGCTATT 3460

RESULT 13
US-10-136-837-3

; Sequence 3, Application US/10136837
; Publication No. US20030181405A1
; GENERAL INFORMATION:
; APPLICANT: NORDSTROM, Jeffrey L
; APPLICANT: PERICLE, Federica
; APPLICANT: ROLLAND, Alain
; APPLICANT: RALSTON, Robert
; TITLE OF INVENTION: Interferon Alpha Plasmids and Delivery Systems, and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 273/082 (GM15-1USC1)
; CURRENT APPLICATION NUMBER: US/10/136,837
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 09/268,135
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5966
; TYPE: DNA

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Mouse IL-12 pin0961
US-10-136-837-3

Query Match 30.3%; Score 1600; DB 13; Length 5966;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 2005; Conservative 0; Mismatches 5; Indels 392; Gaps 1;
QY 2706 CTAGAGCTAATCATGTGCTATAGCTGTTCTCTGTGTGAAATGTTATCCGCTCACAATTC 2765
DB 60 CTTGGCGTAATCATGTGCTATAGCTGTTCTCTGTGTGAAATGTTATCCGCTCACAATTC 119
QY 2766 ACACACATACGAGCGGAGCAATAAGTGTAAAGCTGGGTGCCCTAATGATGAGCTA 2825
DB 120 ACACACATACGAGCGGAGCAATAAGTGTAAAGCTGGGTGCCCTAATGATGAGCTA 179
QY 2826 ACTCACATTAATTTGGCTTGGCTCACTGCGCTCTCTCCAGTCGGGAAACCTGCTGTGCCA 2885
DB 180 ACTCACATTAATTTGGCTTGGCTCACTGCGCTCTCTCCAGTCGGGAAACCTGCTGTGCCA 239
QY 2886 GCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTC 2945
DB 240 GCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTC 299
QY 2946 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
DB 300 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 359
QY 3006 TCACTCAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAAAGAAAT 3065
DB 360 TCACTCAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAAAGAAAT 419
QY 3066 GTGAGCAAAAGCGCAAAAGCCAGGACCGTAAGAGCGGTTGCTGCGCTTTT 3125
DB 420 GTGAGCAAAAGCGCAAAAGCCAGGACCGTAAGAGCGGTTGCTGCGCTTTT 479
QY 3126 CCAATAGGCTCGCGCCCTTGACGAGCATCACAAAATCAACGCTCAAGTCAGAGTGGCG 3185
DB 480 CCAATAGGCTCGCGCCCTTGACGAGCATCACAAAATCAACGCTCAAGTCAGAGTGGCG 539
QY 3186 AAACCCGACAGGACTAAGATACAGGCGTTTCCCTGGAAGCTCCCTCGTGGCTC 3245
DB 540 AAACCCGACAGGACTAAGATACAGGCGTTTCCCTGGAAGCTCCCTCGTGGCTC 599
QY 3246 TCCTGTTCCGACCTCCGCTTACCGATACCTGTCGCTTTCCTTCCGTTCCGGAAGCGT 3305
DB 600 TCCTGTTCCGACCTCCGCTTACCGATACCTGTCGCTTTCCTTCCGTTCCGGAAGCGT 659
QY 3306 GCGCTTTCTCATAGCTACGCTGATAGGTATCTCAGTTCCGTTAGGTGCTTCGCTCCAA 3365
DB 660 GCGCTTTCTCATAGCTACGCTGATAGGTATCTCAGTTCCGTTAGGTGCTTCGCTCCAA 719
QY 3366 GCTGGGCTGTGTCAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAAGTA 3425
DB 720 GCTGGGCTGTGTCAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAAGTA 779
QY 3426 TCCTGTTAGTCAACCCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA 3485
DB 780 TCCTGTTAGTCAACCCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA 839
QY 3486 CAGGATAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAA 3545
DB 840 CAGGATAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAA 899
QY 3546 CTACGGCTACCTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 3605
DB 900 CTACGGCTACCTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 959
QY 3606 CGGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGTGTGGTGGTGGT 3665
DB 960 CGGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGTGTGGTGGTGGT 1019

QY 3666 TTTTGTTCAGCAGCAGGATTAACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGAT 3725
DB 1020 TTTTGTTCAGCAGCAGGATTAACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGAT 1079
QY 3726 CTTTCTTACCGGCTCTGACGCTCAGTGGAAAGAAACTCAGTTAAGGGATTTTGGTCA 3785
DB 1080 CTTTCTTACCGGCTCTGAA----- 1097
QY 3786 GAGATTATCTGCAACAAAGCGGCACTGTCCTCTCCCACTCTCTGCAAGTTTCGGGGCATG 3845
DB 1098 ----- 1097
QY 3846 GATCGCGGATAGCCGCTGCTGTTTCTCGATGCCGACGATTTGCACTGCCGATAGAA 3905
DB 1098 ----- 1097
QY 3906 CTCGCGAGGTGCTCAGCCTCAGGACAGAGCTGAACCAACTCGCGAGGGATCGAGCCC 3965
DB 1098 ----- 1097
QY 3966 GGGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGAGGATCATCCAGCCGGCGTC 4025
DB 1098 ----- 1097
QY 4026 CCGGAAACGATTCGGAAGCCCAACCTTTCATAGAGCGCGGTGGAATCGAAATCTCG 4085
DB 1098 ----- 1097
QY 4086 TGATGGCAGGTTGGCGCTCGCTTGGTGGTCAATTTCCAGACCCAGAGTCCCGCTCAAG 4145
DB 1098 -----CGTCAAG 1107
QY 4146 AACTCGTCAAGAGCGGATAGAGCGATCGCTGCGAATCGGGAGCGGATACCGTAA 4205
DB 1108 AACTCGTCAAGAGCGGATAGAGCGGATCGCTGCGAATCGGGAGCGGATACCGTAA 1167
QY 4206 AGCACGAGGAGCGGTGAGCCCAATTCGCGCCCAAGCTCTTCAGCAATATCACGGTAGCC 4285
DB 1168 AGCACGAGGAGCGGTGAGCCCAATTCGCGCCCAAGCTCTTCAGCAATATCACGGTAGCC 1227
QY 4266 AACGCTATGCTGATAGCGGTGCGCACACCCAGCGCGGACAGTGCATGATCCAGAA 4325
DB 1228 AACGCTATGCTGATAGCGGTGCGCACACCCAGCGCGGACAGTGCATGATCCAGAA 1287
QY 4326 AAGCGGCATTTTCCACATGATATTCGCAAGCAGGATCGCATCGGTGAGTGCAGACAGA 4385
DB 1288 AAGCGGCATTTTCCACATGATATTCGCAAGCAGGATCGCATGCTGCTCAGACAGA 1347
QY 4386 TCCTCGCGCTGGGCAATGCGGCTTGAGCTTGGCGAACAGTTTCGGTGGCGGAGCGCC 4445
DB 1348 TCCTCGCGCTGGGCAATGCGGCTTGAGCTTGGCGAACAGTTTCGGTGGCGGAGCGCC 1407
QY 4446 TGATGCTCTTCGTTCCAGATCATCTGATCGACAAAGACCGGCTTCCATCCGAGTACGTCT 4505
DB 1408 TGATGCTCTTCGTTCCAGATCATCTGATCGACAAAGACCGGCTTCCATCCGAGTACGTCT 1467
QY 4506 CGCTCGATGCGATGTTTTCGCTTGGTGGTGAATGGGAGGTAGCCGATCAAGCGTATGC 4565
DB 1468 CGCTCGATGCGATGTTTTCGCTTGGTGGTGAATGGGAGGTAGCCGATCAAGCGTATGC 1527
QY 4566 AGCGCGCGCATTTGATCAGCATGATGATGATCTTTCGCGAGGAGGAGTGAATGAC 4625
DB 1528 AGCGCGCGCATTTGATCAGCATGATGATGATCTTTCGCGAGGAGGAGTGAATGAC 1587
QY 4626 AGGAGATCTTCGCGCGGCACTTCGCCCAATAGCAGCAGTCCCTTCCCGTTCAGTGACA 4685
DB 1588 AGGAGATCTTCGCGCGGCACTTCGCCCAATAGCAGCAGTCCCTTCCCGTTCAGTGACA 1647
QY 4686 ACCTCGAGCAGGTGGCAAGGAAAGCCCGCTGCTGGCGCAGCCAGATAGCCCGCTGCC 4745
DB 1648 ACCTCGAGCAGGTGGCAAGGAAAGCCCGCTGCTGGCGCAGCCAGATAGCCCGCTGCC 1707
QY 4746 TCGTCTGCGATTTCAATTCAGGGCACCGGACAGGTGCTGCTTGACAAAAAGAACCGGCGC 4805

Db 1708 TCGTCTCGAGTTCAATTCAGGCGACCGGACAGGTTCGTCTTGACAAAAAGAACCGGCGC 1767
Qy 4806 CCCTGCGCTGACACCGGACACACGCGGCGCATCAGAGCAGCCGATTCTGTGTGCGCCAG 4865
Db 1768 CCCTGCGCTGACACCGGACACACGCGGCGCATCAGAGCAGCCGATTCTGTGTGCGCCAG 1827
Qy 4866 TCATAGCGGAATAGCCTCTCCACCCAAAGCGCGGAGAACTCGGTGCAATCATCTTGT 4925
Db 1828 TCATAGCGGAATAGCCTCTCCACCCAAAGCGCGGAGAACTCGGTGCAATCATCTTGT 1887
Qy 4926 TCAATCATGGAACAGATCTCATCTGTCTCTTGATCAGATCTTGATCCCTGCGCCAT 4985
Db 1888 TCAATCATGGAACAGATCTCATCTGTCTCTTGATCAGATCTTGATCCCTGCGCCAT 1947
Qy 4986 CAGATCTTGGCGGCAAGAACGATCCAGTTTACCTTTGAGGCTTCCCAACCTTACCA 5045
Db 1948 CAGATCTTGGCGGCAAGAACGATCCAGTTTACCTTTGAGGCTTCCCAACCTTACCA 2007
Qy 5046 GAGGCGCCCGCAGCTGCAATTCGGTTCGCTTGTCTGCTCCATAAAACCGCCAGTCTAGC 5105
Db 2008 GAGGCGCCCGCAGCTGCAATTCGGTTCGCTTGTCTGCTCCATAAAACCGCCAGTCTAGC 2067
Qy 5106 TA 5107
Db 2068 AA 2069

RESULT 14

US-09-759-960-7/c
; Sequence 7, Application US/09759960
; Patent No. US2001000639A1
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicz, Roman M.
; APPLICANT: Collins, Edward J.
; APPLICANT: Hedley, Mary Lynn
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM THE HPV E7
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/759,960
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/169,425
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/004002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-543-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA

US-09-759-960-7
Query Match 29,7%; Score 1571.6; DB 9; Length 4665;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 114; Indels 180; Gaps 3;
Qy 3008 ACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATTAACGCGAGGAAGAACATGT 3067
Db 2624 AATCATGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGCGAGGAAGAACATGT 2565
Qy 3068 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAAAGGCGCGTGTCTGCGCTTTTCC 3127
Db 2564 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAAAGGCGCGTGTCTGCGCTTTTCC 2505
Qy 3128 ATAGGCTCCGCCCCCTGAGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 3187
Db 2504 ATAGGCTCCGCCCCCTGAGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 2445
Qy 3188 ACCGACAGGACTATAAGATACCAAGGCTTTCCCTCGGAAAGCTCCCTCGTGGCTCTC 3247
Db 2444 ACCGACAGGACTATAAGATACCAAGGCTTTCCCTCGGAAAGCTCCCTCGTGGCTCTC 2385
Qy 3248 CTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAAGCGTG 3307
Db 2384 CTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAAGCGTG 2325
Qy 3308 CGCTTCTCATAGCTCAGCTGATAGGTATCTCAGTTCGGTGTAGGTGGTTCGCTCCAGC 3367
Db 2324 CGCTTCTCATAGCTCAGCTGATAGGTATCTCAGTTCGGTGTAGGTGGTTCGCTCCAGC 2265
Qy 3368 TGGGCTGTGTCAGCAAAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATC 3427
Db 2264 TGGGCTGTGTCAGCAAAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATC 2205
Qy 3428 GTCTTGAAGTCAACCCCGTAAAGACACGACTTATCGCACCTGGGACGACCACTGGTAAAC 3487
Db 2204 GTCTTGAAGTCAACCCCGTAAAGACACGACTTATCGCACCTGGGACGACCACTGGTAAAC 2145
Qy 3488 GGATTAGCAGAGCGAGGTATGTAGCGGTGCTCAGAGTCTTCAAGTGGTGGCTAACT 3547
Db 2144 GGATTAGCAGAGCGAGGTATGTAGCGGTGCTCAGAGTCTTCAAGTGGTGGCTAACT 2085
Qy 3548 ACGGCTACCTAGAGAACAGTATTTGGTATCTCGCTGTCTGCTGAAGCAGATTACCTTCG 3607
Db 2084 ACGGCTACCTAGAGAACAGTATTTGGTATCTCGCTGTCTGCTGAAGCAGATTACCTTCG 2025
Qy 3608 GAAAAGAGTTGGTAGCTCTTGATCCGCAACAAACACCGCTGGTAGCGTGTGTTTT 3667
Db 2024 GAAAAGAGTTGGTAGCTCTTGATCCGCAACAAACACCGCTGGTAGCGTGTGTTTT 1965
Qy 3668 TTGTTTGAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCT 3727
Db 1964 TTGTTTGAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCT 1905
Qy 3728 TTTCTACGGGTCTGACGCTCAGTGGACGAAACTCAGTTAGGGATTTTGGTCTATGA 3787
Db 1904 TTTCTACGGGTCTGACGCTCAGTGGACGAAACTCAGTTAGGGATTTTGGTCTATGA 1845
Qy 3788 GATTATCGTCGACAAAGCGGCTCGTCCCT----- 3819
Db 1844 GATTATCAAAAGGATCTTCACTTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAA 1785
Qy 3820 -----CCCACTCTCGAGTTGCGGG 3840
Db 1784 TCTAAAGTATATATAGTAACCTGAGGCTATGCGAGGCGCTCGCGCCCGACGTTGGCTG 1725
Qy 3841 GCATCGATGCGCGGATAGCCGCTGCTGTTCTTGGATGCGACGAGTTTTCACCTGCGG 3900
Db 1724 CGAGCCCTGGGCTTCACCCGAACTTGGGGGGTGGGGAAGAAAGAACGCGG 1665
Qy 3901 -----TAGAACTCCGCGAGGTCTGTCAGCCTCAGGACGAGTGAACCAAC 3946
Db 1664 CGTATTGGCCCAATGGGCTCTCGTGGGGTATCGACAGAGTGGCCCTGGGACCGAA 1605

Db 3695 GAAAAAGATTGGTAGTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT 3636
QY TTGTTTGCAGACAGCAGATTACGGCGAGAAAAAAGGATCTCAAGAAAGATCTCTTTTGTATCT 3727
Db 3635 TTGTTTGCAGACAGCAGATTACGGCGAGAAAAAAGGATCTCAAGAAAGATCTCTTTGTATCT 3576
QY TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAATGA 3787
Db 3575 TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAATGA 3516
QY GATTATCGTCGACCAAGCGCCATCGTGCT----- 3819
Db 3515 GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAATAAGTTTTAAATCAA 3456
QY 3820 -----CCCACTCTGCACTTCGGGG 3840
Db 3455 TCTAAAGTATATATAGTAACCTGAGGCTATGGCAGGGCTGCGGCCCGACGTTGGCTG 3396
QY 3841 GCATGGATGCGCGATAGCGCTGCTGTTTCTTGGATGCGGACGGATTTGCACTGCGG 3900
Db 3395 CGAGCCCTGGGCTTCCCGGAACTTGGGGGGTGGGGTGGGAAAGAAAGAACCGGG 3336
QY 3901 -----TAGAACTCCCGAGGTCGTCAGCCTCAGGCGAGCAGCTCAACCAAC 3946
Db 3335 GGTATTGCCCCCAATGGGCTCTCGTGGGTAATCGACAGATGCCAGCCCTGGGACCGAA 3276
QY 3947 TCGGAGGGGATCGA----- 3961
Db 3275 CCCCAGCGTTTATGAACAAACGACCCAAACCGTGGTTTTATTCTGTCTTTTATTGGCG 3216
QY 3962 -----GCCCG 3967
Db 3215 TCATAGCGCGGTTCCTTCGGTATTGTCTCCTTCGGTGTTCAGTTAGCTCCCGCTAG 3156
QY 3968 GGTGGGGAAGAAATCCAGCATAGATCCCGCGCTGGAGATCATCCAGCGCGTCCC 4027
Db 3155 GGTGGGGAAGAAATCCAGCATAGATCCCGCGCTGGAGATCATCCAGCGCGTCCC 3096
QY 4028 GGAAGACGATCCGAAGCCCACTTTCATAGAGGGCGGCGGTGGATCGAATCTCGTG 4087
Db 3095 GGAAGACGATCCGAAGCCCACTTTCATAGAGGGCGGCGGTGGATCGAATCTCGTG 3036
QY 4088 ATGCGAGGTTGGGCGTTCGTTGGTCAATTTCCGAACCCAGAGTCCCGCTCAGAGAA 4147
Db 3035 ATGCGAGGTTGGGCGTTCGTTGGTCAATTTCCGAACCCAGAGTCCCGCTCAGAGAA 2976
QY 4148 CTCGTCAAGAGGCGATAGAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTAAAG 4207
Db 2975 CTCGTCAAGAGGCGATAGAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTAAAG 2916
QY 4208 CACGAGGAAGCGGTACGCGCATTCGCGCAAGCTCTTCAGCAATATCAGGGTAGCCAA 4267
Db 2915 CACGAGGAAGCGGTACGCGCATTCGCGCAAGCTCTTCAGCAATATCAGGGTAGCCAA 2856
QY 4268 CGCTATGTCCTGATAGCGGTCCGCCACACCCAGCGCGCCACAGTCGATGAATCCAGAAA 4327
Db 2855 CGCTATGTCCTGATAGCGGTCCGCCACACCCAGCGCGCCACAGTCGATGAATCCAGAAA 2796
QY 4328 GCGGCCATTTTCCACCATGATTTTCGCGNAGCAGGATCGCCATGGGTACAGCAGATC 4387
Db 2795 GCGGCCATTTTTCACCATGATTTTCGCGNAGCAGGATCGCCATGGGTACAGCAGATC 2736
QY 4388 CTCGCGGTGCGGCATGCGCGCTTTGAGCTTGGCGAAACAGTTTCGGCTGGCGGAGCCCTG 4447
Db 2735 CTCGCGGTGCGGCATGCTCGCTTTGAGCTTGGCGAAACAGTTTCGGCTGGCGGAGCCCTG 2676
QY 4448 ATGCTCTTCGTCCAGATCATCTGATCGAAMAGACCGGCTTCCATCCGATACGTGCTCG 4507
Db 2675 ATGCTCTTCGTCCAGATCATCTGATCGAAMAGACCGGCTTCCATCCGATACGTGCTCG 2616
QY 4508 CTCGATGCGATGTTTCGCTTGGTGTGATGGCAGGTAGCGGATCAAGCGGTATGTCAG 4567

Db 2615 CTCGATGCGATGTTTCGCTTGGTGGTCAATGGGCAAGGTAGCCGATCAAGCGGTATGCGAG 2556
QY 4568 CCGCGCATTTGCATCAGCCATGATGATATCTTTCTCGGCAAGGAGCAAGGTGAGATGACAG 4627
Db 2555 CCGCGCATTTGCATCAGCCATGATGATATCTTTCTCGGCAAGGAGCAAGGTGAGATGACAG 2496
QY 4628 GAGATCTTGGCCCCGGCACTTCGCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACAAC 4687
Db 2495 GAGATCTTGGCCCCGGCACTTCGCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACAAC 2436
QY 4688 GTGAGGCAACAGCTGCGCAAGGAACCGCGTCTGGCCAGCCACGATAGCCGCGTGCCTC 4747
Db 2435 GTGAGGCAACAGCTGCGCAAGGAACCGCGTCTGGCCAGCCACGATAGCCGCGTGCCTC 2376
QY 4748 GTCTGCAAGTTCAATTCAGGGGACCCGACAGGTGCGTCTTGACAAAAAGAACCGGGGCGCC 4807
Db 2375 GTCTTGCAGTTCAATTCAGGGGACCCGACAGGTGCGTCTTGACAAAAAGAACCGGGGCGCC 2316
QY 4808 CTGCGCTGACAGCGGAAACACGGCGCATCAGAGCAGCCGATTTGTGTGTGCCAGTC 4867
Db 2315 CTGCGCTGACAGCGGAAACACGGCGCATCAGAGCAGCCGATTTGTGTGTGCCAGTC 2256
QY 4868 ATAGCCGAATAGCCTCTCCACCAGCGGCGCGAGAACCTGCGTGCAATCCATTTGTTTC 4927
Db 2255 ATAGCCGAATAGCCTCTCCACCAGCGGCGCGAGAACCTGCGTGCAATCCATTTGTTTC 2196
QY 4928 AATCATGCGAAACGATCTCTCATCTCTCTTGTGATCAGATCTTG 4971
Db 2195 AATCATGCGAAACGATCTCTCATCTCTCTTGTGATCAGATCTTG 2152

Search completed: February 15, 2004, 10:50:34

Job time : 1750 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 09:54:35 ; Search time 9963 Seconds

(without alignments)
12887.739 Million cell updates/sec

Title: US-09-921-143-36

Perfect score: 5283

Sequence: 1 aagctgaacttatgcgact.....tgagtgcttgaggcagcgtg 5283

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_fod:*
26: em_gss_png:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	991	18.8	2450	11 AK047844	AK047844 Mus muscu
2	948	17.9	1067	9 AU081137	AU081137 AU081137
3	919.8	17.4	1089	9 AU081124	AU081124 AU081124
4	875	16.6	925	14 CB686151	CB686151 Bn01b_02o

5	809.2	15.3	966	29	BZ570738	BZ570738 msh2_1513
6	781.2	14.8	1163	9	AU081044	AU081044 AU081044
7	781.2	14.8	1249	29	BZ572284	BZ572284 msh2_2572
8	780.8	14.8	831	10	BG680919	BG680919 602628716
9	780.8	14.8	858	14	CA488579	CA488579 AGENCOURT
10	780.6	14.8	833	14	CB686421	CB686421 Bn01b_04j
11	774	14.7	954	9	AL044364	AL044364 DKF2p34C
12	754	14.3	1369	29	BZ579291	BZ579291 msh2_6285
13	746	14.1	844	29	BZ574513	BZ574513 msh2_3706
14	742.2	14.0	1346	29	BZ572730	BZ572730 msh2_2771
15	740.8	14.0	914	29	BZ569398	BZ569398 pacs2_164
16	740.4	14.0	831	29	TBR548748	BJ548748 Trypanoso
17	740	14.0	1194	29	BZ578985	BZ578985 msh2_6095
18	734.8	13.9	811	29	ATH517156	AJ517156 Arabidops
19	734.2	13.9	779	9	AL875347	AL875347 AL875347
20	725.6	13.7	959	29	BZ557985	BZ557985 pacs1_60
21	724	13.7	724	9	AL645114	AL645114 AL645114
22	719.6	13.6	847	28	AZ687169	AZ687169 ENTMP05TF
23	719.6	13.6	854	29	BZ570648	BZ570648 msh2_1469
24	718.4	13.6	896	29	BZ551301	BZ551301 pacs1_60
25	715.8	13.5	800	9	AJ281449	AJ281449 4A3A-P4D5
26	715.8	13.5	956	9	AL959357	AL959357 AL959357
27	713.6	13.5	921	11	CNS08VR4	EX030572 Single re
28	711	13.5	1004	9	AJ281480	AJ281480 4A3A-P4G8
29	710.2	13.4	715	9	AL661706	AL661706 AL661706
30	708	13.4	828	29	BZ573391	BZ573391 msh2_3083
31	707.6	13.4	835	29	BZ571572	BZ571572 msh2_1950
32	707	13.4	707	9	AL656688	AL656688 AL656688
33	704	13.3	818	12	B1180988	B1180988 UNL-P-F2-
34	703.4	13.3	705	9	AL635845	AL635845 AL635845
35	703.4	13.3	705	9	AL643164	AL643164 AL643164
36	700.2	13.3	838	28	AZ679307	AZ679307 ENTLG84TF
37	700	13.3	700	13	BQ660293	BQ660293 HT01N05W
38	699.6	13.2	718	9	AL631067	AL631067 AL631067
39	699	13.2	699	13	BQ659852	BQ659852 HF01O22W
40	697.2	13.2	894	29	BZ578323	BZ578323 msh2_5785
41	695.8	13.2	700	13	BQ660226	BQ660226 HT01U04W
42	695	13.2	695	13	BQ660140	BQ660140 HT01C11W
43	695	13.2	697	13	BQ660197	BQ660197 HT01G17W
44	695	13.2	697	13	BQ660261	BQ660261 HT01L03W
45	695	13.2	700	14	CA305674	CA305674 EST000022

ALIGNMENTS

RESULT 1	AK047844	AK047844	2450 bp	linear	HTC 05-DEC-2002
LOCUS	AK047844	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130004M09 product:vascular endothelial growth factor C, full insert sequence.			
DEFINITION	AK047844	AK047844	GI:26339001		
ACCESSION	AK047844	AK047844	GI:26339001		
VERSION	AK047844.1	AK047844.1	GI:26339001		
KEYWORDS	HTC; CAP trapper.	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)	Mus musculus (house mouse)			
ORGANISM	Mus musculus	Mus musculus			
REFERENCE	1	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; 10349636			
AUTHORS	Carninci P. and Hayashizaki, Y.	Carninci P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253	99279253			
PUBMED	10349636	10349636			
REFERENCE	2	10349636			
AUTHORS	Carninci P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Carninci P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374	20499374			
PUBMED	11042159	11042159			

REFERENCE
AUTHORS

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

20530913
11076861
4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaoka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,T., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzairelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2450)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Inotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,P., Takaku-Akai,H.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,I., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1. .2450
/organism="Mus musculus"
/mol_type="rRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C130004M09"
/db_xref="taxon:10090"
/clone="C130004M09"
/issue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
175_1422
/note="unmated protein product; putative
vascular endothelial growth factor C (MGD|GI:109124,
GB|NM_009506, evidence: BLASTN, 99%, match=1784)"
/codon_start=1
/protein_id="BAC33172.1"
/db_xref="GI:26339002"
/translation="WHILCFLSLACLSLLVLPYDWMYKQLKGGWQKQPLNTRIGDS
GGEVKAPEKDELEQLRSVSDVLEMSLVLPYDWMYKQLKGGWQKQPLNTRIGDS
VFAAHAYNTEILKSIDNWRKTQCMREVIDGVKFGAANTFFKPCVSVYRCGG
CNSEGLQMTSTGYLSKLTFTVPLSQGPKVTIIFANHTSCRMKSLDVFQVH
SIIRSLPATLPCQAANKTPTVMNNMCRCLAQDFIYSNVEDDSTNGFDVC
GNKLEDEDTCCVCKGGLRPSGPHKELDRDSCQVCKNKLFPNSCANREFDNT
COCVCKTICPRNQLNPGKCAECENTQKCFKLGKGFHQTCSCVRPNCANRLKHC
PGLSSEEVCRVPSVWKEPHLN"

CDS

BASE COUNT 714 a 541 c 558 g 637 t
ORIGIN
Query Match 18.8%; Score 991; DB 11; Length 2450;
Best Local Similarity 85.4%; Pred. No. 6.1e-266;
Matches 1121; Conservative 0; Mismatches 180; Indels 12; Gaps 1;
QY 481 CACCTGGTGGATCGATCCATCATCGTCTGGGCTTCTCTCTGGCGGTGTTCT 540
Db 151 CTTCCCTGCTCCCGGTCATCCACCATGCTTGTCTTGTCTCTGGCGGTGTTCT 210
QY 541 CTGCTCGCGCTGCGTCTCCCGGTCCTCCGAGGCGCCCGCGCCCGCGCTTC 600
Db 211 CTGCTCGCGCTGCGTCTCCCGGTCCTCCGAGGCGCCCGCGCCCGCGCTTC 270
QY 601 GAGTCCGAGCTCGACTCTTCGAGCGGAGCCCGAGCGGGGTGAGGCCACGGCTATGCA 660
Db 271 GAGTCCGAGCTCGGGCTTCTCGGAAGCGGAGCCCGAGCGGGGTGAGGCCACGGCTTGA 330
QY 661 AGCAAGAGTCTGAGGAGCAGTTCAGTCTGTGTCCAGGTAGTAGTAATCATGCTGTA 720
Db 331 GGCAGAGACTGAGGAGCAGTTCGGCTTGTGTCCAGGTAGTAGTGAATGCTGCTGTC 390
QY 721 CTCTACCCAGATATTGGAAAAATGTACAAGTGTCAAGTCTAGCTAAGGAAAGAGGCTGGCAACAT 780
Db 391 CTGTACCCAGACTACTTGGAAAAATGTACAAGTGTCAAGTGTCAAGTGTCAAGTGTGGC 445
QY 781 AACAGAGACAGGCCAACCTCAACTCAAGGACAGAGAGACTATATAAATTTGCTGCAGCA 840
Db 446 -----AGCAGCCCAACCTCAATACAGGACAGGGGACAGTGTAAAAATTTGCTGCTGCA 498
QY 841 CATTATAACACAGAGATCTTGAAGATATTGATAATAGTGAGGAGAAAGACTCAATGCAATG 900
Db 499 CATTATAACACAGAGATCTTGAAGATATTGATAATAGTGAGGAGAAAGACTCAATGCAATG 558
QY 901 CCACGGAGGTGTGATAGTGTGGGAAGAGGTTTGGAGTCCGACAAAACACCTTCTTT 960
Db 559 CCACGTGAGGTGTGATAGTGTGGGAAGAGGTTTGGAGTCCGACAAAACACCTTCTTT 618
QY 961 AAACCTCCATGTGTGCTGTCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGTGCGAG 1020
Db 619 AAACCTCCATGTGTGCTGTCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGTGCGAG 678
QY 1021 TCGATGAACACACGACGAGCTACTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTC 1080
Db 679 TCGATGAACACACGACGAGCTACTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTC 738
QY 1081 TCTCAAGGCCCCAAACACGATCAATCATCTTTTGCATCATCACTTCTCTCCGATGCAATG 1140

FEATURES	source	Location/Qualifiers
1..1067		
/organism="Oncorhynchus mykiss"		
/mol_type="mRNA"		
/db_xref="taxon:8022"		
/clone="K12"		
/tissue_type="kidney"		
/clone_lib="Oncorhynchus mykiss kidney infected by infectious hematopoietic necrosis virus"		
/note="common name:rainbow trout ; infected by infectious hematopoietic necrosis virus"		
BASE COUNT	235 a	292 c 288 g 250 t
ORIGIN		
Query Match	17.9%	Score 948; DB 9; Length 1067;
Best Local Similarity	98.8%	Pred. No. 4.2e-254;
Matches	976;	Conservative 0; Mismatches 10; Indels 2; Gaps 2;
2706	CTAGACGTAATCATCGTGCATAGCTGTTTCTGCTGTAATTTGTTATCGCTCACAATTC	2765
82	CTTGCGGTAAATCGTGCATAGCTGTTTCTGCTGTAATTTGTTATCGCTCACAATTC	141
2766	ACACAACATAGCAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGCTA	2825
142	ACACAACATAGCAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGCTA	201
2826	ACTCACATTAATGCGTTGCGCTCACTGCCGCTTCCAGTCCGGAAACCTGTCGTGCCA	2885
202	ACTCACATTAATGCGTTGCGCTCACTGCCGCTTCCAGTCCGGAAACCTGTCGTGCCA	261
2886	GCTGCATTAATGAATCGCCAAACGCGGGGAGAGGGGTTTGGTATTTGGGGCTCTTC	2945
262	GCTGCATTAATGAATCGCCAAACGCGGGGAGAGGGGTTTGGTATTTGGGGCTCTTC	321
2946	CGCTTCTCGTCACTGACTGCTGCGCTCGGTGCTTCCGTCGGCGAGCGGTATCAGC	3005
322	CGCTTCTCGTCACTGACTGCTGCGCTCGGTGCTTCCGTCGGCGAGCGGTATCAGC	381
3006	TCACCTCAAAAGCGGTAAATACGGTTATCCACAGATCAGGGGATAAACGACGAAAGACAT	3065
382	TCACCTCAAAAGCGGTAAATACGGTTATCCACAGATCAGGGGATAAACGACGAAAGACAT	441
3066	GTGACAAAGCGCAGCAAAAGGCGAGAACCGTAAAGAGCGCGCTTGTGCGGTTTTT	3125
442	GTGACAAAGCGCAGCAAAAGGCGAGAACCGTAAAGAGCGCGCTTGTGCGGTTTTT	501
3126	CCATAGGCTTCGCGCCCTCGTACGAGCATCAAAAATTCAGCTCAAGTCAGAGTGCGG	3185
502	CCATAGGCTTCGCGCCCTCGTACGAGCATCAAAAATTCAGCTCAAGTCAGAGTGCGG	561
3186	AAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTC	3245
562	AAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTC	621
3246	TCCTGTTCCGACCTCGCGTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGT	3305
622	TCCTGTTCCGACCTCGCGTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGT	691
3306	GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGCTTCCTCGCTCCAA	3365
682	GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGCTTCCTCGCTCCAA	741
3366	GCTGGGCTGTGTGACGAAACCCCGGTTACGCGGACCGTGGCGCTTATCCGGTAACTA	3425
742	GCTGGGCTGTGTGACGAAACCCCGGTTACGCGGACCGTGGCGCTTATCCGGTAACTA	801
3426	TGCTCTTCAGTCCAAACCCCGGTAAACACGACTTATCGCCACTGGCAGCAGCTCGTAA	3485
802	TGCTCTTCAGTCCAAACCCCGGTAAACACGACTTATCGCCACTGGCAGCAGCTCGTAA	861
3486	CAGGATTAAGCAGCGAGGTATGACGCGTGTGTACAGATTTCTGAAGTGGTGGCTAA	3545
862	CAGGATTAAGCAGCGAGGTATGACGCGTGTGTACAGATTTCTGAAGTGGTGGCTAA	921

```

QY 3545 CTACGGCTACACTAGAGAACAGTATTGGTATCTGGCTCTGGCTCTGGAGCCAGTTACCTT 3605
Db 922 CTACGGCTACACTAGAGAACAGTATTGGTATCTGGCTCTGGCTCTGGAGCCAGTTACCTT 981
QY 3606 CGGAAAGAGAGTTGGTAGCTCTTGATCCGGCAACAAACACACCGCTGGTAGCGGTGTTT 3665
Db 982 CGGAAAGAGAGTTGGTAGCTCTTGATCCGGCAACAAACAA-CAACGCTGGTAGCGGTGTTT 1040
QY 3666 TTTTGTTCAGACAGCAGATTACGGC 3693
Db 1041 TTTTGTTC-AGCAGCAGATTACGGC 1067

RESULT 3
AU081124
LOCUS AU081124 1089 bp mRNA linear EST 30-JUL-2002
DEFINITION Oncorhynchus mykiss Kidney infected by infectious
hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG'12,
mRNA sequence.
ACCESSION AU081124.1 GI:6431472
VERSION AU081124.1
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Kono,T., Sakai,M. and Iapetra,S.E.
TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
Location/Qualifiers
FEATURES
source
1..1089
/organism="Oncorhynchus mykiss"
/no_type="mRNA"
/db_xref="taxon:8022"
/clone="KG'12"
/tissue_type="kidney"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/note="common name:rainbow trout ; infected by infectious
hematopoietic necrosis virus"
BASE COUNT 245 a 297 c 295 g 251 t 1 others
ORIGIN

Query Match 17.4%; Score 919.8; DB 9; Length 1089;
Best Local Similarity 97.2%; Pred. No. 3.5e-246;
Matches 979; Conservative 0; Mismatches .22; Indels 6; Gaps 4;

QY 2706 CTAGACGTAATCATGGTCATAGCTTTTCCTGTGTGAATTTATCCGCTCACAATTC 2765
Db 83 CTGCGGTAATCATGGTCATAGCTTTTCCTGTGTGAATTTATCCGCTCACAATTC 142
QY 2766 ACACAACTAGACGCGGAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTA 2825
Db 143 ACACAACTAGACGCGGAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTA 202
QY 2826 ACTCATTAAATGCGTTCGCTCACTCCGCTTTCCAGTCGCGAAACCTGTGCTGCCA 2885
Db 203 ACTCATTAAATGCGTTCGCTCACTCCGCTTTCCAGTCGCGAAACCTGTGCTGCCA 262
QY 2886 GTGCATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 2945
Db 263 GCTGCATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 322

```

```

QY 2946 CGTTTCTCGTCACTGACTGCTCGTTCGCTCGTTCGCTTCGCGGAGCGGTATCAGC 3005
Db 323 CGTTTCTCGTCACTGACTGCTCGTTCGCTCGTTCGCTTCGCGGAGCGGTATCAGC 382
QY 3006 TCACTCAAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACAT 3065
Db 383 TCACTCAAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACAT 442
QY 3066 GTGAGCAAAAGGCGGCAAAAGGCGGAGGAAACCGTAAAAAGGCGGCTTCTCGCGCTTTT 3125
Db 443 GTGAGCAAAAGGCGGCAAAAGGCGGAGGAAACCGTAAAAAGGCGGCTTCTCGCGCTTTT 502
QY 3126 CCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACCTCAAGTCAGAGGTGGCG 3185
Db 503 CCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACCTCAAGTCAGAGGTGGCG 562
QY 3186 AAACCCGACGAGCTATAAGATACGAGCGTTTCCCTGGAAGCTCCCTCGTGGCTC 3245
Db 563 AAACCCGACGAGCTATAAGATACGAGCGTTTCCCTGGAAGCTCCCTCGTGGCTC 622
QY 3246 TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTTCGGAAGCGT 3305
Db 623 TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTTCGGAAGCGT 682
QY 3306 GCGGCTTTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTGTAGGTCTTCGCTCCAA 3365
Db 683 GCGGCTTTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTGTAGGTCTTCGCTCCAA 742
QY 3366 GCTGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAATCA 3425
Db 743 GCTGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAATCA 802
QY 3426 TGCTCTTGTAGTCCAAACCCGCTTACGAGTATTCGCCACTGGCAGGACCACTGGTAA 3485
Db 803 TGCTCTTGTAGTCCAAACCCGCTTACGAGTATTCGCCACTGGCAGGACCACTGGTAA 862
QY 3486 CAGGATTACGAGCGAGGTATGTAGCGGTCTACAGAGTTCTTTGAAG-TGGTGGCCTA 3544
Db 863 CAGGATTACGAGCGAGGTATGTAGCGGTCTACAGAGTTCTTTGAAGTTTGTGGCCTA 922
QY 3545 ACTACGGCTACACTAGAGAACAGTATTGGTATTCGCTGCTGCTGAGCCAGTACT 3604
Db 923 ACTACGGCTACACTAGAGAACAGTATTGGTATTCGCTGCTGCTGAGCCAGTACT 982
QY 3605 TCGGAAAAAGAGTTGGTAGCTCTT-GATCCGCGCAACAAACCA---CCGCTGCTAGCGGT 3660
Db 983 TCGGAAAAAGAGTTGGTAGCTCTTGGATCCGCGCAACAAACCAACACCGCTGGGAGGGGG 1042
QY 3661 GGTTTTTTTTGTGACGCA-CCAGATTACGCGGCAAGAAAAAAGGAT 3706
Db 1043 TGTCTTTTTTTTGTGCAAGCAGCAGATTACCGCGGAAAAAACCGAT 1089

```

RESULT 4

CB86151/c

LOCUS

DEFINITION

Br01b_0208 A

Br01b RAFC ECOR transgenic Brassica napus overexpressing BNCBF17_c

onstitutively_frost_tolerant Brassica napus cDNA clone Br01b_0208,

mRNA sequence.

CB86151

EST.

KEYWORDS

SOURCE

ORGANISM

Brassica napus (rape)

Brassica napus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 925)

REFERENCE

AUTHORS

TITLE

Expressed Sequence Tags from constitutively frost tolerant

transgenic Brassica napus overexpressing BNCBF17

CB86151 925 bp mRNA linear EST 09-APR-2003

Br01b_0208 A

Br01b RAFC ECOR transgenic Brassica napus overexpressing BNCBF17_c

onstitutively_frost_tolerant Brassica napus cDNA clone Br01b_0208,

mRNA sequence.

CB86151

EST.

KEYWORDS

SOURCE

ORGANISM

Brassica napus (rape)

Brassica napus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 925)

REFERENCE

AUTHORS

TITLE

Expressed Sequence Tags from constitutively frost tolerant

transgenic Brassica napus overexpressing BNCBF17

JOURNAL
COMMENT

Unpublished

Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers

FEATURES

source

1. .925
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Westar"
/db_xref="taxon:3708"
/clone="Bn01b_02c08"
/tissue_type="fourth leaf"
/dev_stage="3 weeks seedling grown at room temperature"
/clone_lib="Bn01b_AAPC_ECORC transgenic Brassica napus_ove
repressing ENCBF7 constitutively frost tolerant"
/notes=Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
Site_2: XhoI; Germinated in soil flats and seedlings grown
for 3 weeks in a Conviron E-15 cabinet set at 20°C / 16 hr
light (250 Em-2sec-1) and 16 °C / 8 hr dark. Fourth leaves
collected at 9 am and immediately frozen."
BASE COUNT 218 a 245 c 253 g 203 t
ORIGIN

Query Match 16.6%; Score 875; DB 14; Length 925;
Best Local Similarity 98.6%; Pred. No. 1.2e-233;
Matches 886; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 2701 GCTGTCTAGACGTAATCATGCTCATAGCTGTTTCCTGTGTGAATTTGTTATCGGTCACA 2760
DB GCCTGCTGGGTAATCATAGCTCATAGCTGTTT-CGTGTGAATTTGTTATCGGTCACA 840
QY 2761 ATTCCACACATACGAGCGGAAGCATAAAGTGTAAGCCTCGGGTGCCTTAATGAGTG 2820
DB ATTCCACACATACGAGCGGAAGCATAAAGTGTAAGCCTCGGGTGCCTTAATGAGTG 780
QY 2821 AGTAATCATTAATTTGGTTCGCTCACTGCCGCTTTTCAGTCGGGAACCTGTGC 2880
DB AGTAATCATTAATTTGGTTCGCTCACTGCCGCTTTTCAGTCGGGAACCTGTGC 720
QY 2881 TGCAGCTGATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGTTGTTTGGCGC 2940
DB TGCAGCTGATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGTTGTTTGGCGC 660
QY 2941 TCTTCGCTTCCTCGCTCACTGACTCGTGGCTCGGTTCGTTTCGCTGCGGAGCGGTA 3000
DB TCTTCGCTTCCTCGCTCACTGACTCGTGGCTCGGTTCGTTTCGCTGCGGAGCGGTA 600
QY 3001 TCAGCTCACTCAAGGCGGTAATACGTTTATCCACAGATACAGGAGTAAACGAGGAAG 3060
DB TCAGCTCACTCAAGGCGGTAATACGTTTATCCACAGATACAGGAGTAAACGAGGAAG 540
QY 3061 AACATGTGACAAAGCCAGCAAAAGCGCAGGAACCGTAAAGCCCGCTTGTGCGC 3120
DB AACATGTGACAAAGCCAGCAAAAGCGCAGGAACCGTAAAGCCCGCTTGTGCGC 480
QY 3121 TTTTTCATAGGTCGCGCCCTGACGAGCATCACAAAATTCAGCTCAAGTCAGAGG 3180
DB TTTTTCATAGGTCGCGCCCTGACGAGCATCACAAAATTCAGCTCAAGTCAGAGG 420
QY 3181 TGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTCGTG 3240
DB TGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCTCTGGAAGCTCCCTCGTG 360
QY 3241 CGCTCTCTGTCGACCTGCGCTTACCGGATACCTGTGCGCTTTTCTCCCTTCGGA 3300
DB CGCTCTCTGTCGACCTGCGCTTACCGGATACCTGTGCGCTTTTCTCCCTTCGGA 300
QY 3301 AGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTGCTGTAGTTCGCTCGC 3360

Db 299 AGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTGCTGTAGTTCGTCGC 240
QY 3361 TCCAGCTGGCTGTGTGTCAGCAACCCCGCTTCAGCCCGACCGCTCGCTTATCGGT 3420
Db 239 TCCAGCTGGCTGTGTGTCAGCAACCCCGCTTCAGCCCGACCGCTCGCTTATCGGT 180
QY 3421 AACTATCGTCTTGTAGTCCAAACCCCGTAAAGACAGACTTATCGCCACTGGCAGCACT 3480
Db 179 AACTATCGTCTTGTAGTCCAAACCCCGTAAAGACAGACTTATCGCCACTGGCAGCACT 120
QY 3481 GGTAAACGATTAACAGAGGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTG 3540
Db 119 GGTAAACGATTAACAGAGGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTG 60
QY 3541 CCTAACTACGGCTACACTAGAAACAGTATTTGGTATCTGCGCTCTGCTGAAACCACT 3599
Db 59 CTAACTACGGCTACACTAGAAACAGTATTTGGTATCTGCGCTCTGCTGAAACCACT 1

RESULT 5
B2570738
LOCUS
DEFINITION B2570738 966 bp DNA linear GSS 17-DEC-2002
msh2_1513.x1 msh Pseudomonas aeruginosa genomic clone msh2_1513,
genomic survey sequence.
ACCESSION B2570738
VERSION B2570738.1 GI:27205799
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 966)
REFERENCE
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol., (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun
FEATURES
source
1. .966
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_1513"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."
BASE COUNT 223 a 267 c 266 g 209 t
ORIGIN

Query Match 15.3%; Score 809.2; DB 29; Length 966;
Best Local Similarity 97.0%; Pred. No. 3.7e-215;
Matches 845; Conservative 0; Mismatches 24; Indels 2; Gaps 2;

QY 2890 CATTAAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCT 2949
Db 94 CATTAAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCT 153
QY 2950 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3009
Db 154 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 213
QY 3010 TCAAAGCGGTAAATACGGTTATCCAGAAATCAGGGGATAACGAGGAAGAAACATGTGA 3069

Db 214 TCAAGGCGGTATATACGGTTTATCCACAGAAATACAGGGATACAGGAGAAAGACATGTGA 273
QY 3070 GCAAAAGCCAGCAAAAGCCAGAGAACCGTAAAGAGCCGGTTGCTGGGGTTTTTCCAT 3129
Db 274 GCAAAAGCCAGCAAAAGCCAGAGAACCGTAAAGAGCCGGTTGCTGGGGTTTTTCCAT 333
QY 3130 AGGCTCCGCGCCCTTGACGAGATCACAAAAATCGAGCTCAAGTCAAGAGGTGGCGAAAC 3189
Db 334 AGGCTCCGCGCCCTTGACGAGATCACAAAAATCGAGCTCAAGTCAAGAGGTGGCGAAAC 393
QY 3190 CGCAGGAGCTATTAAGATACAGCGGTTTCCCTCGAGAGCTCCCTCGTGGCTTCCT 3249
Db 394 CGCAGGAGCTATTAAGATACAGCGGTTTCCCTCGAGAGCTCCCTCGTGGCTTCCT 453
QY 3250 GTTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCCTTCGGGAAGCGTGGCG 3309
Db 454 GTTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCCTTCGGGAAGCGTGGCG 513
QY 3310 CTTTCTCATAGCTACGCTGAGTATCTCAGTTTCGGTGTAGGTCGTTCCGCTCCAGCTG 3369
Db 514 CTTTCTCATAGCTACGCTGAGTATCTCAGTTTCGGTGTAGGTCGTTCCGCTCCAGCTG 573
QY 3370 GGCTGTGTGACGAAACCCCGCTTACGCGGAGCTGCGCTTATCCGGTAACTATCGT 3429
Db 574 GGCTGTGTGACGAAACCCCGCTTACGCGGAGCTGCGCTTATCCGGTAACTATCGT 633
QY 3430 CTTGAGTCCAAACCCGGTAAGACAGACTTATCGGCACTGGCAGAGCACTGTAAACAGG 3489
Db 634 CTTGAGTCCAAACCCGGTAAGACAGACTTATCGGCACTGGCAGAGCACTGTAAACAGG 693
QY 3490 ATTACAGAGGAGTATGTAGGGGCTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTAC 3549
Db 694 NATACAGAGGAGTATGTAGGGGCTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTAC 753
QY 3550 GGCTACACTAGAAAGATATTTGGTATCTGGCTCTGTGAAGCAGATTAACCTCGGA 3609
Db 754 GGCTACACTAGAAAGATATTTGGTATCTGGCTCTGTGAAGCAGATTAACCTCGGA 812
QY 3610 AAAAGAGTGTGTAGCTTTGATCGCGCAAAACAAACCCGCTGTAGCGGTGGTTTTTT 3669
Db 813 AAAAGAGTGTGTAGCTTTGATCGCGCAAAACAAACCCGCTGTAGCGGTGGTTTTTT 872
QY 3670 GTTTCGACGAGAGATTAACGCGCAGAAAGAAAGATCTCAAGAGATTCCTTGAATCTT 3729
Db 873 GTTTCGACGAGCAAAATACGCGCAGAAAGAAAGATCTCAAGAGATTCCTTGAATCTT 932
QY 3730 TCTAGGGGCTGACGCTCAGTGGAACGAA 3760
Db 933 T-TACGGGTTTCGAGCTCAAGGAGCCGAAA 962

RESULT 6
LOCUS AU081044
DEFINITION AU081044.1 linear EST 30-JUL-2002
hematopoietic necrosis virus Oncorhynchus mykiss cdna clone KBI,
mRNA sequence.
ACCESSION AU081044
VERSION AU081044.1 GI:6431392
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 1163)
Kono, T., Sakai, M. and Lapetra, S.E.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
Contact: Masahiro Sakai
Faculty of Agriculture

Miyazaki University
1-1 nishi gakukenbanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
Location/Qualifiers
1. 1163
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="KBI"
/tissue_type="kidney"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/note="common name:rainbow trout ; infected by infectious
hematopoietic necrosis virus"
BASE COUNT 272 a 307 c 298 g 286 t
ORIGIN

Query Match 14.8%; Score 781.2; DB 9; Length 1163;
Best Local Similarity 99.5%; Pred. No. 3.1e-207;
Matches 794; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 2706 CTAGACGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATTC 2765
Db 367 CTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATTC 426
QY 2766 ACACACATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGCTA 2825
Db 427 ACACACATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAGCTA 486
QY 2826 ACTCACAATTAATTTGGCTTTCGCTCACTCCCGCTTTCCAGTGGGAAACCTGTCTGTGCA 2885
Db 487 ACTCACAATTAATTTGGCTTTCGCTCACTCCCGCTTTCCAGTGGGAAACCTGTCTGTGCA 546
QY 2886 GTTCGATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 2945
Db 547 GTTCGATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 606
QY 2946 CGCTTCCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
Db 607 CGCTTCCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 666
QY 3006 TCACCTCAAGAGCGGTAAATACGTTATCCACAGAAATCAGGGATACAGGAGAAACAT 3065
Db 667 TCACCTCAAGAGCGGTAAATACGTTATCCACAGAAATCAGGGATACAGGAGAAACAT 726
QY 3066 GTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCA 3125
Db 727 GTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCA 786
QY 3126 CCATAGGCTCCGCGCCCTCGAGCAGATCACAAAAATCGAGTCAAGTCAAGTCAAGTCAAG 3185
Db 787 CCATAGGCTCCGCGCCCTCGAGCAGATCACAAAAATCGAGTCAAGTCAAGTCAAGTCAAG 846
QY 3186 AAACCCGACAGACTATTAAGATACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTC 3245
Db 847 AAACCCGACAGACTATTAAGATACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTC 906
QY 3246 TCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCCTTCGGGAAGCGT 3305
Db 907 TCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCCTTCGGGAAGCGT 966
QY 3306 GCGCTTTCTCATAGCTCAGCTGAGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCCAA 3365
Db 967 GCGCTTTCTCATAGCTCAGCTGAGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCCAA 1026
QY 3366 GTTGGGCTGTGTGACAGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACTA 3425
Db 1027 GCTGGGCTGTGTGACAGAA-CCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACTA 1085
QY 3426 TGTCTTGTAGTCAACCCCGGTAAAGACAGACTTATCGCCACTGGCAGAGCCACTGTAA 3485
Db 1086 TGTCTTGTAGTCAACCCCGGTAAAGACAGACTTATCGCCACTGGCAGAGCCACTGTAA 1145


```

QY 3486 CAGATTAGCAGACGAG 3503
Db 1146 CAGATTAGCAGACGAG 1163

RESULT 7
BZ572284 1249 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_2572.x3 msh Pseudomonas aeruginosa genomic clone msh2_2572,
DEFINITION genomic survey sequence.
ACCESSION BZ572284
VERSION BZ572284.1 GI:27207345
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
AUTHORS Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1249)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol., (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
    source
        1..1249
            Location/Qualifiers
                /organism="Pseudomonas aeruginosa"
                /mol_type="genomic DNA"
                /strain="MSH"
                /db_xref="taxon:287"
                /clone_lib="msh2_2572"
                /notes="Environmental isolate. Whole genomic shotgun
                    library."
BASE COUNT 292 a 337 c 323 g 294 t 3 others
ORIGIN

Query Match 14.8%; Score 781.2; DB 29; Length 1249;
Best Local Similarity 95.4%; Pred. No. 3.2e-207;
Matches 857; Conservative 0; Mismatches 35; Indels 6; Gaps 5;

QY 2883 CCAGCTGCATTATGAATCGCCCAACGCGCGGAGAGCGGTTTCGGTATTGGCGCTC 2942
Db 100 CCAGCTGCATTATGAATCGCCCAACGCGCGGAGAGCGGTTTCGGTATTGGCGCTC 159

QY 2943 TTCGGCTTCCTCGCTCACTGACTCGCTGCGCTCGCTCGCTTCGGCTCGCGCGAGCGGTATC 3002
Db 160 TTCGGCTTCCTCGCTCACTGACTCGCTGCGCTCGCTTCGGCTCGCGCGAGCGGTATC 219

QY 3003 AGTCTACTCAAGGCGGTAATACGGTTATCCAGAAATCAGGGGATAACGAGGAAGAA 3062
Db 220 AGTCTACTCAAGGCGGTAATACGGTTATCCAGAAATCAGGGGATAACGAGGAAGAA 279

QY 3063 CATGTGAGCAAAAGGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGCGCTTCGTCGCGTT 3122
Db 280 CATGTGAGCAAAAGGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGCGCTTCGTCGCGTT 339

QY 3123 TTTCATAGCTCCGCGCCCTCGAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTG 3182
Db 340 TTTCATAGCTCCGCGCCCTCGAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTG 399

QY 3183 GCGAAACCCGACAGGACTATAAGATACGAGGCGTTTCCCTCGGAGCTCCCTCGTCG 3242
Db 400 GCGAAACCCGACAGGACTATAAGATACGAGGCGTTTCCCTCGGAGCTCCCTCGTCG 459

QY 3243 CTCTCCTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAG 3302

```

```

Db 460 CTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAG 519
QY 3303 CGTGGCGCTTCTCATAGCTCAGCTCAGCTGTAGTATCTCAGTTCGGTGTAGGTTCGTC 3362
Db 520 CGTGGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTGTAGGTTCGTC 579
QY 3363 CAAGTGGCGCTGTGTGACAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAA 3422
Db 580 CAAGTGGCGCTGTGTGACAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAA 639
QY 3423 CTATCGCTCTTGAGTCCAAACCGGTAAGACACGACTTATGCGCACTGCGCAGCAGCACTGG 3482
Db 640 CTATCGCTCTTGAGTCCAAACCGGTAAGACACGACTTATGCGCACTGCGCAGCAGCACTGG 699
QY 3483 TAA CAGGATTAGCAGACGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGCC 3542
Db 700 TAA CAGGATTAGCAGACGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGCC 759
QY 3543 TAACTACGGCTACACTAGAAACAGTATTTGGTATCTGCGTCTGCTGCTGAAGCAGTTAC 3602
Db 760 TAACTACGGCTACACTAGAAACAGTATTTGGTATCTGCGTCTGCTGCTGAAGCAGTTAC 818
QY 3603 CTTCGAAAAAGAGTTGGTAGCTCTTGATCGGCAAAACAAACCAACCGCTGTAGCGGTGG 3652
Db 819 CTTCGAAAAAGAGTT-GTAGCTCTTGATCG--CAACAAACCAACCGCTGTAGCGGTGG 875
QY 3663 TTTTGTGTTTGGCAAGCAGAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTT 3722
Db 876 -NTTTTGGTTGCAAGCAGCA-ATTACGCGCAGAAAAAGGTTCTCAGGAAGATCTTT 933
QY 3723 GATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAACCTCAGTTAAGGATTTTG 3780
Db 934 GTTCTTTCTACGGGCTGACGCTTGTGGCCCTTAGTGGAGCAAACTCAGTAAGGATTTTG 991

```

```

RESULT 8
BG680919 602628716F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753755 5',
LOCUS mRNA sequence.
ACCESSION BG680919.1 GI:13912303
VERSION BG680919
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 831)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10613 row: 0 column: 04
High quality sequence stop: 810.
Location/Qualifiers
    1..831
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4753755"
        /tissue_type="squamous cell carcinoma"
        /lab_host="DH10B (TI phage-resistant)"
        /clone_lib="NCI CGAP Skn4"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;

```

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
205 c 214 g 182 t

BASE COUNT	230 a	205 c	214 g	182 t
ORIGIN				
Query Match	14.8%; Score 780.8; DB 10; Length 831;			
Best Local Similarity	99.3%; Pred. No. 3.2e-207;			
Matches 826; Conservative	0; Mismatches 2; Indels 4; Gaps 4;			
QY	536	GTTCTTCGCTCGCCGGCTCGCTGTCTCCGGGTCTCGCAGAGCGGCCGCCGCCGCCGC	595	
DB	1	GTTCCTCTCTCGCCGCTCGCTGTCTCCGGGTCTCGCAGAGCGGCCGCCGCCGCCGC	60	
QY	596	CCTTCGAGTCCGGACTCGAATCTCTCGACCGAGGCCGACGCGGTGAGGCCACCGGTT	655	
DB	61	CCTTCGAGTCCGGACTCGAATCTCTCGACCGAGGCCGACGCGGTGAGGCCACCGGTT	120	
QY	656	ATGCAAGCAAAAGATCTGGAGGAGCAGATTACGGTCTGTTCAGTGTAGATGAACATCA	715	
DB	121	ATGCAAGCAAAAGATCTGGAGGAGCAGATTACGGTCTGTTCAGTGTAGATGAACATCA	180	
QY	716	CTGTACTCTACCCAGAATATTGAAAAATGTACAAGTGTCAAGTAAGGAAGAGGCTGCG	775	
DB	181	CTGTACTCTACCCAGAATATTGAAAAATGTACAAGTGTCAAGTAAGGAAGAGGCTGCG	240	
QY	776	AACATAACAGAGAACAGGCCAACCTCAACTCTAAGGACAGAGAGACTATAAAATTGCTG	835	
DB	241	AACATAACAGAGAACAGGCCAACCTCAACTCTAAGGACAGAGAGACTATAAAATTGCTG	300	
QY	836	CAGCACATTAATACAGAGATCTTGAAAAGTATTGATATGAGTGGAGAAAGACTCAAT	895	
DB	301	CAGCACATTAATACAGAGATCTTGAAAAGTATTGATATGAGTGGAGAAAGACTCAAT	360	
QY	896	GCATGCCACGGGAGGTGTATAGATGTGGGGAGAGAGTTTGAGTTCGGACGCAACACCT	955	
DB	361	GCATGCCACGGGAGGTGTATAGATGTGGGGAGAGAG-TTGAGTTCGGACGCAACACCT	419	
QY	956	TCTTTAAACCTCCATGTGTTCCTCTACAGATGTGGGGGTGCTGCAATAGTAGGGGC	1015	
DB	420	TCTTTAAACCTCCATGTGTTCCTCTACAGATGTGGGGGTGCTGCAATAGTAGGGGC	479	
QY	1016	TGCAGTGCATGAACACCAGCAGCAGGTACCTCAGCAAGAGCTTATTTTGAAATTCAGTGC	1075	
DB	480	TGCAGTGCATGAACACCAGCAGCAGGTACCTCAGCAAGAGCTTATTTTGAAATTCAGTGC	539	
QY	1076	CTCTCTCTCAAGGCCCAAACACGATPACAATCATGTTTTGCCAATCACATCTCTCCCGAT	1135	
DB	540	CTCTCTCTCAAGGCCCAAACACGATPACAATCATGTTTTGCCAATCACATCTCTCCCGAT	599	
QY	1136	GCATGTCTAACTGGATGTTTACAGACAGTTCAATTCATTTAGACGTCCTCCGCGCAG	1195	
DB	600	GCATGTCTAACTGGATGTTTACAGACAGTTCAATTCATTTAGACGTCCTCCGCGCAG	659	
QY	1196	CAACACTACCAAGTGTCCAGGACGGAACAGAGCTGCCCCCAACCAATTACATGTGGAATA	1255	
DB	660	CAACA-TACCAAGTGTCCAGGACGGAACAGAGCTGCCCCCAACCAATTACATGTGGAATA	718	
QY	1256	ATCACATCTGCAGATGCTGCTCAGGAAGATTATTATGTTTTCTCGGATGCTGCAGATG	1315	
DB	719	ATCACATCTGCAGATGCTGCTCAGGAAG-TTATATGTTTTCTCGGATGCTGCAGATG	777	
QY	1316	ACTCAACAGATGGATTCCATGCATCTGTGGACCAACAGGAGCTGGATGA	1367	
DB	778	ACTCAACAGATGGATT-CATGCATCTGTGGACCAACAGGAGCTGGATGA	828	

RESULT	9
CA488579	
LOCUS	CA488579
DEFINITION	AGENCOURT_10808789 MAPCJ Homo sapiens cDNA clone IMAGE:6720351 5', mRNA sequence.
ACCESSION	CA488579
	858 bp mRNA linear EST 14-NOV-2002

VERSION CA488579.1 GI:24950807
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE	1 (bases 1 to 858)
AUTHORS	NIH-MGC http://mgc.ncbi.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rcapbs@mail.nih.gov Tissue Procurement: Kristi A. Eglund, Ira Pastan cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L14M14279 row: 1 column: 15 High quality sequence stop: 674.

FEATURES

```

1. 858
/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="taxon:9606"
/clone_image="6720351"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL1"
' LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPCL"
/notes=vector; pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Straussberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
257 a 205 C 199 G 195 t 2 others

```

Query Match	14.8%; Score 780.8; DB 14; Length 858;
Best Local Similarity	99.0%; Pred.No.3.3e-207;
Matches 816; Conservative 0; Mismatches 4; Indels 3;	
QY	883 AGAAGACTCAATGCATGCCACGGGAGGTCGTATAGATGTGGGGAAGAGCTTTGGAGTC 942
DB	1 AGAAGACTCAATGCATGCCACGGGAGGTCGTATAGATGTGGGGAAGAGCTTTGGAGTC 60
QY	943 GCGCAAAACACCTTCTTTAAACCTCCATGTGTCCGTCTACAGATGTGGGGGTGTGTCG 1002
DB	61 GCGCAAAACACCTTCTTTAAACCTCCATGTGTCCGTCTACAGATGTGGGGGTGTGTCG 120
QY	1003 AATAGTGTGGGGGTGCAGTGCATGAACACGACGAGCTACCTCAGCAAGACGCTTATTT 1062
DB	121 AATAGTGTGGGGGTGCAGTGCATGAACACGACGAGCTACCTCAGCAAGACGCTTATTT 180
QY	1063 GAAATTCAGTGCCTCTCTCAAGGCCCCAAACAGATAACAATCAGTTTTGGCAATCAC 1122
DB	181 GAAATTCAGTGCCTCTCTCAAGGCCCCAAACAGATAACAATCAGTTTTGGCAATCAC 240
QY	1123 ACTTCTCCGCGATGCATGTCTAAACTGGATGTTTACAGCAAGTTCATTCATTATTAGA 1182
DB	241 ACTTCTCCGCGATGCATGTCTAAACTGGATGTTTACAGCAAGTTCATTCATTATTAGA 300
QY	1183 CGTTCCTCCGCGAGCAACATACACAGTGTCAAGGACGGCAACAGACCTGCCCCACCAAT 1242
DB	301 CGTTCCTCCGCGAGCAACATACACAGTGTCAAGGACGGCAACAGACCTGCCCCACCAAT 360
QY	1243 TACATGTGGGAATAATCACATCTGCAGATGCCTGGCTCAGGAAGATTTTATGTTTCTCG 1302
DB	361 TACATGTGGGAATAATCACATCTGCAGATGCCTGGCTCAGGAAGATTTTATGTTTCTCG 420

QY 1303 GATCTCGAGTACTCAACAGATGGATTCCATGACATCTGTGGACCAACAGAGCTG 1362
 Db 421 GATCTCGAGTACTCAACAGATGGATTCCATGACATCTGTGGACCAACAGAGCTG 480
 QY 1363 GATCAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTTCGAGCTGTGGACCC 1422
 Db 481 GATCAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTTCGAGCTGTGGACCC 540
 QY 1423 CACAAAGAACTAGACAGAACTCATGCCAGTGTCTGTAAACAACTCTTCCCGAGC 1482
 Db 541 CACAAAGAACTAGACAGAACTCATGCCAGTGTCTGTAAACAACTCTTCCCGAGC 600
 QY 1483 CAATGTGGGCGCAACCGAGAACTTGTATGAACACATGCGAGTGTGTATGAAAGAAC 1542
 Db 601 CAATGTGGGCGCAACCGAGAACTTGTATGAACACATGCGAGTGTGTATGAAAGAAC 660
 QY 1543 TGCCCCAGAAATCAACCCCTTAAATCTCGGAAATGTGCTGTGAATGTACAGAAAGTCCA 1602
 Db 661 TGCCCCAGAAATCAACCCCTTAAATCTCGGAAATGTGCTGTGAATGTACAGAAAGTCCA 720
 QY 1603 CAGAAATGCTGTGTTAAAGGAAAGAGTTCACACCAACCAACATGCGAGTGTACAG-ACG 1661
 Db 721 CAGAAATGCTGTGTTAAAGGAAAGAGTTCACACCAACCAACATGCGAGTGTACAGAAAG 780
 QY 1662 GCCATGTACG- AACCGCGAGAGGC--TTGTGAGCGAGATTTT 1702
 Db 781 GCCATGTACGAAACCGCGAGAGGCCTTTGTGAGCGAGATTTT 824

RESULT 10
 CB686421/c
 LOCUS
 DEFINITION
 Bn01b_04j14.4 833 bp mRNA linear EST 09-APR-2003
 Bn01b_AFC_ECORC transgenic Brassica napus overexpressing BNCBF17 c
 onstitutively frost tolerant Brassica napus cDNA clone Bn01b_04j14,
 mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica napus (rape)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 1; eurosids II; Brassicales; Brassicaceae; Brassica.
 Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moors, A.,
 Chagnon, J., Fatah, S., Couroux, P., and Hattori, J.
 Expressed Sequence Tags from constitutively frost tolerant
 transgenic Brassica napus overexpressing BNCBF17
 Unpublished
 Contact: Singh, J.A.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
 0C6, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singhja@agr.ca.
 Location/Qualifiers

1
 833
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Westar"
 /db_xref="taxon:3708"
 /clone="Bn01b_04j14"
 /tissue_type="fourth leaf"
 /dev_stage="3 weeks seedling grown at room temperature"
 /clone_lib="Bn01b_AFC_ECORC transgenic Brassica napus_ove
 repressing BNCBF17 constitutively frost tolerant"
 /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
 Site 2: XhoI; Germinated in soil flats and seedlings grown
 for 3 weeks in a Conviron E-15 cabinet set at 20°C/16 hr

FEATURES

source

BASE COUNT 189 a 223 c 238 g 182 t 1 others
 ORIGIN
 Query Match
 Best Local Similarity 14.8%; Score 780.6; DB 14; Length 833;
 Matches 805; Conservative 1; Mismatches 5; Indels 2; Gaps 2;
 QY 2701 GCTGCTAGACGTAATCATGCTGCTATAGCTGTTTCCTGTGTGAATTTATTCGGTCA 2760
 Db 812 GCGCGCTGGCGTAATCATGCTGCTATAGCTGTTTCCTGTGTGAATTTATTCGGTCA 753
 QY 2761 ATTCACACAACTACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTTAATGAGTG 2820
 Db 752 ATTCACACAACTACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTTAATGAGTG 693
 QY 2821 AGCTAACTCACTTAATTTGGTTGGCTCACTCGCGCTTTCCAGTCCGGGAAACCTGTGCG 2880
 Db 692 AGCTAACTCACTTAATTTGGTTGGCTCACTCGCGCTTTCCAGTCCGGGAAACCTGTGCG 633
 QY 2881 TGCAGCTGCAATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTCCGCTATTGGGGCGC 2940
 Db 632 TGCAGCTGCAATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTCCGCTATTGGGGCGC 573
 QY 2941 TCTTCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 3000
 Db 572 TCTTCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 513
 QY 3001 TCAGCTCACTCAAGGGGTATACGGTTATCCAGAGATCAGGGGATAACGCGAGGAAG 3060
 Db 512 TCAGCTCACTCAAGGGGTATACGGTTATCCAGAGATCAGGGGATAACGCGAGGAAG 453
 QY 3061 AACATGTGAGCAAAAGGCGCAAAAGGCGGAGAAACCGTAAAGAGGCGCGTTCTGGCG 3120
 Db 452 AACATGTGAGCAAAAGGCGCAAAAGGCGGAGAAACCGTAAAGAGGCGCGTTCTGGCG 393
 QY 3121 TTTTTCATAGGCTCGCGCCCTCGAGAGCATCAAAATTCGACGCTCAAGTCAGAGG 3180
 Db 392 TTTTTCATAGGCTCGCGCCCTCGAGAGCATCAAAATTCGACGCTCAAGTCAGAGG 333
 QY 3181 TGGGAAACCGGACGAGACTATAAGATACAGGGGTTTCCCGCTGGAGCTCCCTCGTG 3240
 Db 332 TGGGAAACCGGACGAGACTATAAGATACAGGGGTTTCCCGCTGGAGCTCCCTCGTG 273
 QY 3241 CGCTCTCTCTTCGAGCCCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGA 3300
 Db 272 CGCTCTCTCTTCGAGCCCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGA 213
 QY 3301 AGCGTGGCGCTTCTCATAGCTCAAGCTGATGATCTCAGTTGGGTAGTGGTTCG 3360
 Db 212 AGCGTGGCGCTTCTCATAGCTCAAGCTGATGATCTCAGTTGGGTAGTGGTTCG 153
 QY 3361 TCCAAGTGGCGTGTGTGCAACCCCGCTTCAGCCCGGAGCGCTGCCTTAT-CCGG 3419
 Db 152 TCCAAGTGGCGTGTGTGCAACCCCGCTTCAGCCCGGAGCGCTGCCTTAT-CCGG 93
 QY 3420 TAATCTATGCTTGTAGTCCAAACCGGTGAAGACACGACTTATCGGCACCTGGCAGAGCCAC 3479
 Db 92 TAATCTATGCTTGTAGTCCAAACCGGTGAAGACACGACTTATCGGCACCTGGCAGAG-CAC 34
 QY 3480 TGGTAAAGATTTAGCAGAGCGAGGTATGTAGG 3512
 Db 33 TGGTAAAGATTTAGCAGAGCGAGGTATGTAGG 1

RESULT 11

AL044364

LOCUS

DEFINITION

ACCESSION

VERSION

AL044364 954 bp mRNA linear EST 29-FEB-2000
 DXFP434C172 s1.434 (synonym: htes3) Homo sapiens cDNA clone
 DXFP434C172 s1.434, mRNA sequence.
 AL044364
 AL044364.1 GI:5432586

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 954)
 AUTHORS Ansoerge.W., Benes.V., Krieger.S., Mewes.H.W., Gassenhuber,J. and Wiemann.S.
 TITLE EST (Ansoerge, Benes, et al.)
 JOURNAL Unpublished
 COMMENT Contact: Ansoerge W
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 3' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the CDNA sequencing consortium of the German Genome Project.
 No r1 sequence available.
 This clone (DKFZp343C172) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source
 1..954
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp343C172"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="vector: pSPort1; Site_1: Not1; Site_2: Sal1"
 BASE COUNT 214 a 273 c 253 g 214 t
 ORIGIN
 Query Match 14.7%; Score 774; DB 9; Length 954;
 Best Local Similarity 100.0%; Pred. No. 2.8e-205;
 Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2887 CTGCATTAAATGAAATCGGCCAACCGCGGGGAGAGCGGTTGGGTATGGCGCTTTCC 2946
 Db 181 CTGCATTAAATGAAATCGGCCAACCGCGGGGAGAGCGGTTGGGTATGGCGCTTTCC 240
 2947 GCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCT 3006
 Db 241 GCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCT 300
 3007 CACTCAAAGCGGTAATACGTTATCCACAGAAATCAGGGGATACCGAGAAAGACATG 3066
 Db 301 CACTCAAAGCGGTAATACGTTATCCACAGAAATCAGGGGATACCGAGAAAGACATG 360
 3067 TGAGCAAAAGCCAGCAAAAGCCAGCAAGCCGTAAGGCGCGGTTGCTGGCGTTTTC 3126
 Db 361 TGAGCAAAAGCCAGCAAAAGCCAGCAAGCCGTAAGGCGCGGTTGCTGGCGTTTTC 420
 3127 CATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGGCGA 3186
 Db 421 CATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGGCGA 480
 3187 AACCCGACGAGCTATAAGATACGAGCGGTTCCCGTGAAGCTCCCTCGTGGCTCT 3246
 Db 481 AACCCGACGAGCTATAAGATACGAGCGGTTCCCGTGAAGCTCCCTCGTGGCTCT 540
 3247 CTTGTTCCGACCTCGCCCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 3306
 Db 541 CTTGTTCCGACCTCGCCCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG 600
 3307 GCGCTTTCTCAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGTCCCAAG 3366
 Db 601 GCGCTTTCTCAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGTCCCAAG 660

3367 CTGGGCTGTGTGCACGAAACCCCGCTTACGCCGAGCGCTGGCGCTTATCCGGTAACTAT 3426
 Db 661 CTGGGCTGTGTGCACGAAACCCCGCTTACGCCGAGCGCTGGCGCTTATCCGGTAACTAT 720
 3427 CGTCTTGGTCCAAACCCGGTAAAGACAGCACTTATGCCACTGGCAGCAGCACTGGTAAC 3486
 Db 721 CGTCTTGGTCCAAACCCGGTAAAGACAGCACTTATGCCACTGGCAGCAGCACTGGTAAC 780
 3487 AGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAAC 3546
 Db 781 AGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAAC 840
 3547 TAGGGCTACACTAGAAACAGATTTTGGTATCTGGCTCTGCTGAAGCAGTTACTCTTC 3606
 Db 841 TAGGGCTACACTAGAAACAGATTTTGGTATCTGGCTCTGCTGAAGCAGTTACTCTTC 900
 3607 GGAATAAGAGTTGGTAGCTCTTGCATCCGCAAAACCAACCCGCTGGTAGCGGT 3660
 Db 901 GGAATAAGAGTTGGTAGCTCTTGCATCCGCAAAACCAACCCGCTGGTAGCGGT 954
 RESULT 12
 BZ579291 1369 bp DNA linear GSS 17-DEC-2002
 LOCUS msh2_6285.x1 msh Pseudomonas aeruginosa genomic clone msh2_6285,
 DEFINITION genomic survey sequence.
 ACCESSION BZ579291
 VERSION BZ579291.1 GI:27214352
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 REFERENCE 1 (bases 1 to 1369)
 AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.
 TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
 JOURNAL J. Bacteriol., (2002) In press
 COMMENT Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers
 1..1369
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="MSH"
 /db_xref="taxon:287"
 /clone="msh2_6285"
 /clone_lib="msh"
 /note="Environmental isolate. Whole genomic shotgun library."
 BASE COUNT 327 a 356 c 351 g 334 t 1 others
 ORIGIN
 Query Match 14.3%; Score 754; DB 29; Length 1369;
 Best Local Similarity 95.6%; Pred. NO. 1.5e-199;
 Matches 797; Conservative 0; Mismatches 35; Indels 2; Gaps 2;
 2840 CGTTGGGCTCACTGCCCGCTTCCAGTCGGAAACCTGTCGTCGAGTGCATTAATGAA 2899
 Db 99 CGTTGGGCTCACTGCCCGCTTCCAGTCGGAAACCTGTCGTCGAGTGCATTAATGAA 158
 2900 TCGGGCCAAACGCGGGGAGAGCGGTTTCGGTATTGGCGGCTCTTCCGCTTCTCGCTCA 2959
 Db 159 TCGGGCCAAACGCGGGGAGAGCGGTTTCGGTATTGGCGGCTCTTCCGCTTCTCGCTCA 218
 2960 CTGACTCGTTCGGCTCGGTCGCTTCGGCTCGCGGAGCGGTATCAGTCACTCAAGGCGG 3019

```

Db 219 CTGACTCGCTCGCTCGGTGTTTCGGCTGGCGAGCGGTATCAGTCACTCAAGAGCGG 278
Qy 3020 TAATACAGTTATCCACAGATCAGGGATAACGAGGAAAGAAACATGTGAGCAAAAGGCC 3079
Db 279 TAATACAGTTATCCACAGATCAGGGATAACGAGGAAAGAAACATGTGAGCAAAAGGCC 338
Qy 3080 AGCAAAAGGCCAGGAAACGTAAGAAAGCCCGCTTGTGGCTTTTCCATAGAGCTCCGCC 3139
Db 339 AGCAAAAGGCCAGGAAACGTAAGAAAGCCCGCTTGTGGCTTTTCCATAGAGCTCCGCC 398
Qy 3140 CCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAGGAC 3199
Db 399 CCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAGGAC 458
Qy 3200 TATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGGAGCGGTGCGCTTCTCATATA 3259
Db 459 TATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGGAGCGGTGCGCTTCTCATATA 518
Qy 3260 TGGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAGCGGTGCGCTTCTCATATA 3319
Db 519 TGGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAGCGGTGCGCTTCTCATATA 578
Qy 3320 GCTCACGCTGTAGTATCTCAGTTCCGCTGAGTGTGCTTCCCTCCAAAGCTGGGCTGTGTGC 3379
Db 579 GCTCACGCTGTAGTATCTCAGTTCCGCTGAGTGTGCTTCCCTCCAAAGCTGGGCTGTGTGC 638
Qy 3380 ACGNACCCCGCTTACGCGCGAGCGCTGCGCTTATCCCGTAACATCGTCTTGAAGTCCA 3439
Db 639 ACGNACCCCGCTTACGCGCGAGCGCTGCGCTTATCCCGTAACATCGTCTTGAAGTCCA 698
Qy 3440 ACCCGGTAAAGACAGCACTTATCGCACTGCGCAGCAGCACTGCTTAACAGATTTAGCAGAG 3499
Db 699 ACCCGGTAAAGACAGCACTTATCGCACTGCGCAGCAGCACTGCTTAACAGATTTAGCAGAG 758
Qy 3500 CGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGTGGCGCTAACTACGCTACACTA 3559
Db 759 CGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGTGGCGCTAACTACGCTACACTA 818
Qy 3560 GAAGAAC-AGTATTGTTATCTCGCTCTGCTGAGCCAGTACCTTCGGAAGAGAT 3618
Db 819 GAAGAACAGATTTTGGTTTGGCTTCTGCTGAGCCAGTACCTTTGGAAGAGAT 878
Qy 3619 GGTAGCTCTTGATCCGGCAAAACCAACCGC-TGGTAGCGGTGTTTTTTTGT 3671
Db 879 GGTAGCTCTTGATCCGGCAAAACCAACCGCGGTGGTAGCGGTGTTTTTTT 932

```

RESULT 13

BZ574513

LOCUS

DEFINITION

msh2_3706.x1 msh Pseudomonas aeruginosa genomic clone msh2_3706,

genomic survey sequence.

ACCESSION

BZ574513

VERSION

BZ574513.1 GI:27209574

KEYWORDS

GSS.

SOURCE

Pseudomonas aeruginosa

Pseudomonas aeruginosa

Pseudomonadaceae; Pseudomonas.

REFERENCE

AUTHORS

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,

Burns, J.L., Kaul, R. and Olsen, M.V.

TITLE

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol., (2002) In press

COMMENT

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

FEATURES

source

Location/Qualifiers

1..844

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="msh"

/db_xref="taxon:287"

/clones="msh2_3706"

/clone_lib="msh"

/note="Environmental isolate. Whole genomic shotgun

library."

BASE COUNT

202 a 225 c 222 g 193 t 2 others

ORIGIN

Query Match

Best Local Similarity

Matches 749; Conservative

14.1%; Score 746; DB 29; Length 844;

99.1%; Pred. No. 1.9e-197;

0; Mismatches 7; Indels 0; Gaps 0;

Qy 3039 ATCAGGGATTAACGAGGAAAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCG 3098

Db 87 ATCAGGGATTAACGAGGAAAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCG 146

Qy 3099 TAAAGAGCCCGCTTGTGGCTTTTCCATAGAGTCCGCCCTCCAGAGCATCAAA 3158

Db 147 TAAAGAGCCCGCTTGTGGCTTTTCCATAGAGTCCGCCCTCCAGAGCATCAAA 206

Qy 3159 AAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAAGATACCAAGCGTT 3218

Db 207 AAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAAGATACCAAGCGTT 266

Qy 3219 TCCCGCTGGAAGTCCCTCGTGGCTTCCAGAGTCCGCTTCCAGAGTCCGCTTCCAGAGTCCCT 3278

Db 267 TCCCGCTGGAAGTCCCTCGTGGCTTCCAGAGTCCGCTTCCAGAGTCCGCTTCCAGAGTCCCT 326

Qy 3279 GTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGTCACTGATAGTATCT 3338

Db 327 GTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGTCACTGATAGTATCT 386

Qy 3339 CAGTTCGGTGTAGTGTGCTCCAGCTGGGCTGTGTGACAGCAAAACCCCGCTTCAGCC 3398

Db 387 CAGTTCGGTGTAGTGTGCTCCAGCTGGGCTGTGTGACAGCAAAACCCCGCTTCAGCC 446

Qy 3399 CGACGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGGTAAGACACGACTT 3458

Db 447 CGACGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGGTAAGACACGACTT 506

Qy 3459 ATCGGCACTGCGAGCAGCACTGTGTAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGC 3518

Db 507 ATCGGCACTGCGAGCAGCACTGTGTAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGC 566

Qy 3519 TACAGAGTTCTTGAAGTGTGGCTTAACCTAGCGCTACACTAGAGAAACAGTATTTGGTAT 3578

Db 567 TACAGAGTTCTTGAAGTGTGGCTTAACCTAGCGCTACACTAGAGAAACAGTATTTGGTAT 626

Qy 3579 CTGCGCTGTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGTCTTGTATCCGCCAA 3638

Db 627 CTGCGCTGTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGTCTTGTATCCGCCAA 686

Qy 3639 ACAAGCAGCCGCTGTAGCGGTGTTTTTTTGTTCGAGCAGAGTATTCGCGCAAAA 3698

Db 687 ACAAGCAGCCGCTGTAGCGGTGTTTTTTTGTTCGAGCAGAGTATTCGCGCAAAA 746

Qy 3699 AAAAGGATCTCAAGAGATCTTTTGTATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGA 3758

Db 747 AAAAGGATCTCAAGAGATCTTTTGTATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGA 806

Qy 3759 AAATCAGTTAAGGATTTTGGTCAATGATATTC 3794

Db 807 NAATCAGTTAAGGATTTTGGTCAATGATATTC 842

RESULT 14

BZ572730

```

LOCUS       B2572730                1346 bp    DNA    linear    GSS 17-DEC-2002
DEFINITION  msh2_2771.x1 msh Pseudomonas aeruginosa genomic clone msh2_2771,
            genomic survey sequence.
ACCESSION   B2572730
VERSION     B2572730.1  GI:27207791
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
            Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1346)
            Spencer D.H., Raymond C.K., Smith E.E., Sims E.E., Hastings M.,
            Burns J.L., Kaul R. and Olsen M.V.
            Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
            J. Bacteriol., (2002) In press
            Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 20622216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..1346
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="MSH"
                     /db_xref="taxon:287"
                     /clone="msh2_2771"
                     /clone_lib="msh"
                     /notes="Environmental isolate. Whole genomic shotgun
                     library."
BASE COUNT      299 a    369 c    346 g    330 t    2 others
ORIGIN

```

Query Match	14.0%	Score 742.2	DB 29	Length 1346
Best Local Similarity	87.9%	Pred. No. 3e-196		
Matches 855	Conservative 0	Mismatches 110	Indels 8	Gaps 4
QY	2835	AAATGCGTTGGCTCACTGCGCGCTTCCAGTGGGAAACCTGTCGGTGCAGTGCATTA	2894	
DB	84	AAITCTGTCAGCCCACTGCTGTTTCCAGTGGGAAA	143	
QY	2895	ATGAATCGGCCAACCGCGCGGGAGAGCGGTTTCGTTATTTGGGCGCTCTTTCCGCTTCCTC	2954	
DB	144	ATGAATCGGCCAACCGCGCGGGAGAGCGGTTTCGTTATTTGGGCGCTCTTTCCGCTTCCTC	203	
QY	2955	GCTCACTGACTTCGCTGCGCTCGGTCGTTTCGCTCGCGCGAGCGGTATCAGTCACTCAAA	3014	
DB	204	GCTCACTGACTTCGCTGCGCTCGGTCGTTTCGCTCGCGCGAGCGGTATCAGTCACTCAAA	263	
QY	3015	GGCGGTTAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAAA	3074	
DB	264	GGCGGTTAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAAA	323	
QY	3075	AGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCCGCGTTCGTCGCGCTTTTCCATAGGCT	3134	
DB	324	AGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCCGCGTTCGTCGCGCTTTTCCATAGGCT	383	
QY	3135	CCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC	3194	
DB	384	CCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC	443	
QY	3195	AGGACTATAAGATACCAAGGCGTTTCCCGCTGGAAGCTCCCTCGTCGCGCTCTCTCTGTTCC	3254	
DB	444	AGGACTATAAGATACCAAGGCGTTTCCCGCTGGAAGCTCCCTCGTCGCGCTCTCTCTGTTCC	503	
QY	3255	GACCTGCGGTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAGCGTGGCGCTTTC	3314	
DB	504	GACCTGCGGTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAGCGTGGCGCTTTC	563	

Qy	3315	TCATAGCTCAGCGTGTAGGTTATCTCAGTTCCGGTGTAAGTTCGGTCCGAAGTCGGCGTG	3374
Dd	564	TCATAGCTCAGCGTGTAGGTTATCTCAGTTCCGGTGTAAGTTCGGTCCGAAGTCGGCGTG	623
Qy	3375	TGTGCACGAACCCCCGTTTCAGCCCAGCCGCTCGGCCCTTATCCGGTAAACTATCGTCTTTGA	3434
Dd	624	TGTGCACGAACCCNCGTTCCGCCCAGCCGCTCGGCCCTTATCCGGTAAACTATCGTCTTTGA	683
Qy	3435	GTTCAACCCTGGTTAAGCACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACA - GGATT	3493
Dd	684	GTTCAACCCTGGTTAAGCACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACA	743
Qy	3494	GCAGCGGAGGTATGTAGCGCGTCTACAGAGTTCCTTAGTGGTGGCCTTAACACCGCT	3553
Dd	744	GCAGCGGAGGTATGTAGCGCGTCTACAGAGTTCCTTAGTGGTGGCCTTAACACCGCT	803
Qy	3554	ACACTAGAAGAACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGGTTACCTTCGGA	3613
Dd	804	TACTTGAAGAACCGATTGGTAACTGCGCTCTGGTGAACCCGTTACCTTT - GGA	861
Qy	3614	GAGTTGGTAGCTCTTGATCCGGGCAACAAACACCCGCTGGTAGCGGTGTTTTGTTT	3673
Dd	962	GAAAGTGGTAGCTTTTGATCCGG - AAAACAAACCCACGCTGGTGC	920
Qy	3674	GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTCATCTTTCTA	3733
Dd	921	GCAGCCCAATTTACGCG - ----AAAATAAAGACTTCAGAAAACTTTGATCTTTCAC	976
Qy	3734	CGGGGTCTGAGCTCATGTGAACGAAACTCAGCTTAAGGATTTGGTCATGAGATTAT	3793
Dd	977	CGGGTGTCTCCAGGAGCGAAATCCACGTTAGGTTTTGTGATTAGATATCAAAGGT	1036
Qy	3794	CGTCGACCAAGC	3806
Dd	1037	CTTCCCTCAATC	1049
RESULT	15	BZ569398	
LOCUS		pacs2-164_8370.xl pacs2-164_Pseudomonas aeruginosa genomic clone	
DEFINITION		pacs2-164_8370, genomic survey sequence.	
ACCESSION		BZ569398	
VERSION		BZ569398.1	
KEYWORDS		GI:27203968	
SOURCE		GSS.	
ORGANISM		Pseudomonas aeruginosa	
REFERENCE		Pseudomonas aeruginosa	
AUTHORS		Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;	
		Pseudomonadaceae; Pseudomonas.	
TITLE		1 (bases 1 to 914)	
JOURNAL		Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,	
COMMENT		Burns,J.L., Kaul,R. and Olsen,M.V.	
		Whole-Genome-Sequence variation among multiple isolates of	
		Pseudomonas aeruginosa library	
		J. Bacteriol., (2002) In press	
		Contact: Chris K. Raymond	
		Genome Center	
		University of Washington	
		Box 352145, Seattle, WA 98105-2145, USA	
		Tel: 20622216954	
		Fax: 2066857244	
		Email: craymond@u.washington.edu	
		Class: shotgun.	
FEATURES		Location/Qualifiers	
source		1..914	
		/organism="Pseudomonas aeruginosa"	
		/mol_type="genomic DNA"	
		/strain="2-164"	
		/db_xref="taxon:287"	
		/clone="pacs2-164_8370"	
		/clone_lib="pacs2-164"	
		/notes="clinical isolate 2-164 Whole genomic shotgun	
		library."	

BASE COUNT		201 a	260 c	238 g	206 t	9 others	
ORIGIN							
Query Match		14.0%; Score 740.8; DB 29; Length 914;					
Best Local Similarity		98.7%; Pred. No. 5.8e-196;					
Matches 756; Conservative		0; Mismatches 9; Indels 1; Gaps 1;					
QY	2938	CGCTCTTCGGCTTCCTCGCTCAGTCACTCGCTGCGCTCGGTCGTTGCGCTGCGCGAGCG					2997
Db	110	CGCTCTTCGGCTTCCTCGCTCAGTCACTCGCTGCGCTCGGTCGTTGCGCTGCGCGAGCG					169
QY	2998	GTATCAGCTCACTCAAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGA					3057
Db	170	GTATCAGCTCACTCAAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGA					229
QY	3058	AAGACATGTGAGCAAAAGGCGGCAAAAGGCGGCAAGGCGGCGGTTGCTG					3117
Db	230	AAGACATGTGAGCAAAAGGCGGCAAAAGGCGGCAAGGCGGCGGTTGCTG					289
QY	3118	GCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAG					3177
Db	290	GCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAG					349
QY	3178	AGGTGGCGAACC CGACAGACTATAAAGATACCAGGCGTTTCCCTCGAAGCTCCCTC					3237
Db	350	AGGTGGCGAACC CGACAGACTATAAAGATACCAGGCGTTTCCCTCGAAGCTCCCTC					409
QY	3238	GTGGCTCTCTGTTTCCGACCTCGCGCTTACCGGATACCTGTCGGCTTCTCCCTTCG					3297
Db	410	GTGGCTCTCTGTTTCCGACCTCGCGCTTACCGGATACCTGTCGGCTTCTCCCTTCG					469
QY	3298	GGAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTT					3357
Db	470	GGAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTT					529
QY	3358	CGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTGCGCTTATCC					3417
Db	530	CGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTGCGCTTATCC					589
QY	3418	GGTAACATCTCTTTGAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGGCAGAGCC					3477
Db	590	GGTAACATCTCTTTGAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGGCAGAGCC					649
QY	3478	ACTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTCAAGTGG					3537
Db	650	ACTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTCAAGTGG					709
QY	3538	TGGCCTTAACGCTACACTAGAGAACAGTATTTGGTATCTGCGCTCTCTGCTGAAGCCA					3597
Db	710	TGGCCTTAACGCTACACTAGAGAACAGTATTTGGTATCTGCGCTCTCTGCTGAAGCCA					769
QY	3598	GTTACCTTCGGAAGAGTTGGTGTAGTCTTTGATCCGGCAACAAACACCGCTGGTAGC					3657
Db	770	GTTACCTTCGGAAGAGTTGGTGTAGTCTTTGATCCGGCAACAAACACCGCTGGTAGC					828
QY	3658	GGTGGTTTTTTTGTGCAAGCAGAGATTACGCGCAGAAAAAAG					3703
Db	829	GGTGGTTTTTTTGTGCAAGCAGAGATTACGCGCAGAAAAAAG					874

Search completed: February 15, 2004, 19:07:05
Job time : 9974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:10:26 ; Search time 1979.02 Seconds
(without alignments)
10335.846 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500

Sequence: 1 aagctgaccttgatgact.....cacctgggtggatgcacgc 500

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_ov.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	36.4	1560	14	AELENV K02374 Rous-associ
2	175	35.0	9396	14	AF052428 AF052428 Rous earc
3	140	28.0	346	14	AELETR2 K00993 Rous-associ
4	127	25.4	281	6	BD136066 BD136066 Novel pro
5	127	25.4	282	6	BD136065 BD136065 Novel pro
6	127	25.4	512	6	BD136064 BD136064 Novel pro
7	127	25.4	590	6	E59416 E59416 Signal pept
8	127	25.4	599	6	E51990 E51990 Promoter, r
9	127	25.4	600	6	AX360741 AX360741 Sequence
10	127	25.4	619	6	AX256412 AX256412 Sequence
11	127	25.4	738	6	AX591955 AX591955 Sequence
12	127	25.4	777	6	BD000145 BD000145 Vector ha
13	127	25.4	777	6	I05430 I05430 Sequence 16
14	127	25.4	777	6	I08105 I08105 Sequence 3
15	127	25.4	875	14	HS5PTE1AB M64941 Human cytom
16	127	25.4	876	14	HS5PTE1AE M64944 Human cytom
17	127	25.4	876	14	HS5PTE1AF M64940 Human cytom
18	127	25.4	877	14	HS5PTE1AC M64942 Human cytom
19	127	25.4	919	6	BD000143 BD000143 Vector ha
20	127	25.4	919	6	I05393 I05393 Sequence 14
21	127	25.4	919	6	I08103 I08103 Sequence 1
22	127	25.4	930	6	A10617 A10617 Human Cytom
23	127	25.4	930	6	AR028792 AR028792 Sequence
24	127	25.4	930	6	AR028794 AR028794 Sequence
25	127	25.4	930	6	AR050544 AR050544 Sequence
26	127	25.4	930	6	AR050546 AR050546 Sequence
27	127	25.4	930	6	AR094363 AR094363 Sequence
28	127	25.4	930	6	AR241121 AR241121 Sequence
29	127	25.4	930	6	AR241123 AR241123 Sequence
30	127	25.4	930	6	BD131785 BD131785 Transfect
31	127	25.4	930	6	BD145110 BD145110 Methods a
32	127	25.4	930	6	BD145112 BD145112 Methods a
33	127	25.4	930	6	E00836 E00836 Recombinant
34	127	25.4	930	6	I49834 I49834 Sequence 2
35	127	25.4	930	6	I49836 I49836 Sequence 4
36	127	25.4	930	14	HS5IEE K03104 Human Cytom
37	127	25.4	996	6	AX282811 AX282811 Sequence
38	127	25.4	1070	6	A85308 A85308 Sequence 6
39	127	25.4	1070	6	BD107647 BD107647 FIV vacci
40	127	25.4	1078	6	A92081 A92081 Sequence 5
41	127	25.4	1078	6	AR230688 AR230688 Sequence
42	127	25.4	1318	6	A92080 A92080 Sequence 4
43	127	25.4	1318	6	A9230690 A9230690 Sequence
44	127	25.4	1417	6	A92077 A92077 Sequence 1
45	127	25.4	1417	6	AR230692 AR230692 Sequence

ALIGNMENTS

RESULT 1

AELENV

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AELENV

Rous-associated virus-2 env and gag genes.

K02374

env protein; gag protein; long terminal repeat (LTR).

Rous-associated virus type 2

Rous-associated virus type 2

Viruses; Retroviral viruses; Retroviridae; Alpharetrovirus.

1 (bases 1 to 1560)

Bizub,D., Katz,R.A. and Skalka,A.M.

Nucleotide sequence of noncoding regions in Rous-associated

virus-2: comparisons delineate conserved regions important in

replication and oncogenesis

1560 bp ss-RNA linear

VRL 28-APR-1993


```
mat_peptide      2130..2501
/gene="gag"
/product="p15"
<2522..5210
/gene="pol"
<2522..5209
/gene="pol"
/codon_start=1
/product="polymerase and integrase protein"
/protein_id="AAC08988.1"
/db_xref="GI:3003002"
/translation="TVALHAIPLKWKPDHTFPWIDOWLPPEGLKVALTQVKELOL
GHLEPSLQNTWTFVIRKASGSRLLHDLRAVNAKLVFPFQVQAGPVLSPALRGPW
LWLDLQDFFSFLAEQDREAFPLPSVNOAPARRQFWKVLPGMGTSPCTICQLV
VGVLEPLRLKPSLRMLHMDLLAAASHDGLAAGEVINTLERAGFTTSPKIQ
REGVOYLGLKGSTVAVPGLVAEPRIATLDVOKLVGSLQWLPALGIPRLMGPF
YEQLRSDPNEAREWNLDMKMAWEIVQLSTTAALERMDPALPLEGAVRCEQGLAV
LGGDLFPLPGIILALRGFAGKINSSTPSIFDIARPLVSLKRVTDHPVPTFTD
FREDLPLPGIILALRGFAGKINSSTPSIFDIARPLVSLKRVTDHPVPTFTD
ASSSTHGVVWVRGPRWEIKETADIGASVOOLEARAVAMALLMPTTTPVNTVSAP
VAKMLKMGQGVPTAAAFILDALQSASMAAVLHVRSHEVPGFTFGNDVADSQ
ATFOAYLREAKLHTTLHIGPRALSKACNISMQOAREVVOTCPHCNAPALEAGVNP
RGLGPIQWTDFTLSPRMAPESLWAVTDTASSAIVTQHGRTVSVAQGHWATAIA
VLGRPKAKIDNGSCSTSKSTREWLARWCIATHTGIPGNSQOQAVERNRLKXKIR
VLAEGGFMKRIPASKQGEKLLAKAMTALNHFERTGNTKTPVQKHWRPVTLTGPPVKI
RIETGWEKGMNVLVWGRYAAVKNRDTDKVWPSRVKVPDITQKDEVTKKDEASPL
FAGSDWIPWGDEQGLQEEAASNKQEGFCEDTLAANES"
/repeat_region 6961..7040
/rpt_type=direct
7300..8908
/gene="src"
7300..8907
/gene="src"
/codon_start=1
/product="src tyrosine kinase"
/protein_id="AAC08990.1"
/db_xref="GI:3003004"
/translation="MSSSKSKPKDPHRRSRLEPPDSTHGGFPASQTPNKTAAPDTH
RTPSRFGTAVTEKFLGDFNTDVTSPQRGALAGGVTTFVALDYDSWLETDLSP
KGERLQIVNNTGNNWLAHSVTGCTGIVPSNYAPSDSIQAEWYFGKIRFRESER
LLNPNRPTFLRWSLETTKGAYGTSVDFDNAGLVNKHVKIRKLDGGGYITTSRT
QFSLQLOVAYYSKHADGLCHRLTNVCPTSPQOTGLAKDAWEIPRESLRLEKVLQV
CFGEVWMTGNTGTGTGTGTVAIKTLKPGTSPPEAPLOBAOVAKLREKVLQV
YAVVSEPIVLVLEYSKSLDPLKGMGKVLRLPOLDMAAQIASGMAYVERNNYV
HDLRAANILVGENLVCKVDFGLKLEEDNEYTARQCAKFIKWTAPAAIYGRFTI
KSDVWFGILLTETLTKGRYPFGMNGEVLDRVGRYKMPCPPECPSLHDLNMQCCW
RRDPERPFTFEXLQALLPACVLEVAE"
/repeat_region 9002..9096
/rpt_type=direct
9131..9396
9131..9375
/note="U3"
/polyA_signal 9369..9374
/repeat_region 9376..9396
/note="R"
BASE COUNT      2226 a 2366 c 2740 g 2064 t
ORIGIN
35.0%; Score 175; DB 14; Length 9396;
Query Match
Best Local Similarity 100.0%; Pred. No. 5e-95;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      259 TGAGTTAGCAACATGCCCTTATAGGAGAGAAAAAGACCGTGCATGCCGATTGGTGGGAG 318
Db      9183 TGAGTTAGCAACATGCCCTTATAGGAGAGAAAAAGACCGTGCATGCCGATTGGTGGGAG 9242
QY      319 TAAGTGGTATGATCGTGGTATGATCGTGCCTTGTAGGAAGCGCAACAGACGGGTCTAAC 378
Db      9243 TAAGTGGTATGATCGTGGTATGATCGTGCCTTGTAGGAAGCGCAACAGACGGGTCTAAC 9302
QY      379 ACGGATTGGACGACCACTGAATTCGCGATTGCAGAGATATTCGTTAAGTGC 433
Db      9303 ACGGATTGGACGACCACTGAATTCGCGATTGCAGAGATATTCGTTAAGTGC 9357
```

```
RESULT 3
ALELTR2
LOCUS
DEFINITION
endonuclease cleavage sites.
ACCESSION
K00993
VERSION
GI:210075
KEYWORDS
cleavage site; endonuclease cleavage site; long terminal repeat
(LTR); reverse transcriptase endonuclease cleavage site.
SOURCE
Rous-associated virus type 2
ORGANISM
Rous-associated virus type 2
VIRUSES; Retrovirdae; Retroviridae; Alpharetrovirus.
REFERENCE
1 (bases 1 to 346)
Duyk, G., Leis, J., Longiaru, M. and Skalka, A.M.
Selection cleavage in the avian retroviral long terminal repeat
sequence by the endonuclease associated with the alpha beta form of
avian reverse transcriptase
Proc. Natl. Acad. Sci. U.S.A. 80 (22), 6745-6749 (1983)
84070706
MEDLINE
6196775
PUBMED
2 (bases 1 to 346)
Skalka, A.M., Duyk, G., Longiaru, M., DeHaseth, P., Terry, R. and
Leis, J.
Integrative recombination--a role for the retroviral reverse
transcriptase
Cold Spring Harb. Symp. Quant. Biol. 49, 651-659 (1984)
85153042
MEDLINE
6085046
PUBMED
COMMENT
Original source text: Rous-associated virus 2 replication form I
DNA, clone RAV2-2 [1], and plasmids pP01 and pGJ14 [2].
Reverse transcriptase associated endonuclease (purified from avian
sarcoma virus) cleavage sites have been mapped in two tandemly
linked Rous-associated virus-2 LTR sequences. The enzyme may be
involved in viral cDNA integration in the host, since it generates
a 6 bp staggered overlap that spans the junction. The clone
sequence (RAV2-2) corresponds to the unintegrated replicative form
(RF) I of RAV-2.
Draft entry for [2] kindly provided by A. Skalka, 15-AUG-1985.
Location/Qualifiers
1..346
/organism="Rous-associated virus type 2"
/mol_type="genomic RNA"
/db_xref="taxon:11948"
<1..154
/note="LTR A"
complement(104..105)
/note="rt-endonuclease cleavage site"
130..131
/note="rt-endonuclease secondary cleavage site"
151..152
/note="rt-endonuclease primary cleavage site"
155..1546
/note="LTR B"
complement(157..158)
/note="rt-endonuclease primary cleavage site"
214..215
/note="rt-endonuclease secondary cleavage site"
complement(222..223)
/note="rt-endonuclease secondary cleavage site"
complement(244..245)
/note="rt-endonuclease secondary cleavage site"
304..305
/note="rt-endonuclease secondary cleavage site"
BASE COUNT      96 a 69 c 89 g 92 t
ORIGIN
EcORI site.
Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e-73;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      259 TGAGTTAGCAACATGCCCTTATAGGAGAGAAAAAGACCGTGCATGCCGATTGGTGGGAG 318
```

```

Db      207 TGAGTTAGCAATCCCTTATAAGAGAGAAAAAGACCGTGCATGCCGATTGGTGGAG 266
QY      319 TAAAGTGGTATGATCGTGTATGATCGCTTGTAGGAGGCAACAGACGGGTCTAAC 378
Db      267 TAAAGTGGTATGATCGTGTATGATCGCTTGTAGGAGGCAACAGACGGGTCTAAC 326
QY      379 ACGGATTGACGAACCACTG 398
Db      327 ACGGATTGACGAACCACTG 346

RESULT 4
LOCUS   BD136066
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136066
VERSION   BD136066.1 GI:23231011
KEYWORDS JP 2002508974-A/3.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
1 (bases 1 to 281)
Armentano,D., Yew,N. and Marshall,J.
Novel promoter element for sustained gene expression
Patent: JP 2002508974-A 3 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/3
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression
Location/Qualifiers
FT source 1..281
FT /organism='Cytomegalovirus'

Query Match 25.4%; Score 127; DB 6; Length 281;
Best Local Similarity 100.0%; Pred.No.1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 76
Db 31 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 90
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 136
Db 91 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 150
QY 137 CACCCCA 143
Db 151 CACCCCA 157

RESULT 5
LOCUS   BD136065
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136065
VERSION   BD136065.1 GI:23231010
KEYWORDS JP 2002508974-A/2.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
1 (bases 1 to 281)
Armentano,D., Yew,N. and Marshall,J.
Novel promoter element for sustained gene expression
Patent: JP 2002508974-A 3 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/2
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression
Location/Qualifiers
FT source 1..281
FT /organism='Cytomegalovirus'

Query Match 25.4%; Score 127; DB 6; Length 281;
Best Local Similarity 100.0%; Pred.No.1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 76
Db 31 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 90
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 136
Db 91 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 150
QY 137 CACCCCA 143
Db 151 CACCCCA 157

```

```

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
1 (bases 1 to 282)
Armentano,D., Yew,N. and Marshall,J.
Novel promoter element for sustained gene expression
Patent: JP 2002508974-A 2 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/2
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression
Location/Qualifiers
FT source 1..282
FT /organism='Cytomegalovirus'

Query Match 25.4%; Score 127; DB 6; Length 282;
Best Local Similarity 100.0%; Pred.No.1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 76
Db 27 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 86
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 136
Db 87 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 146
QY 137 CACCCCA 143
Db 147 CACCCCA 153

RESULT 6
LOCUS   BD136064
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136064
VERSION   BD136064.1 GI:23231009
KEYWORDS JP 2002508974-A/1.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
1 (bases 1 to 512)
Armentano,D., Yew,N. and Marshall,J.
Novel promoter element for sustained gene expression
Patent: JP 2002508974-A 1 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/1
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression
Location/Qualifiers
FT source 1..512
FT /organism='Cytomegalovirus'

Query Match 25.4%; Score 127; DB 6; Length 282;
Best Local Similarity 100.0%; Pred.No.1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 76
Db 27 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 86
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 136
Db 87 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 146
QY 137 CACCCCA 143
Db 147 CACCCCA 153

RESULT 6
LOCUS   BD136064
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136064
VERSION   BD136064.1 GI:23231009
KEYWORDS JP 2002508974-A/1.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
1 (bases 1 to 512)
Armentano,D., Yew,N. and Marshall,J.
Novel promoter element for sustained gene expression
Patent: JP 2002508974-A 1 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/1
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression
Location/Qualifiers
FT source 1..512
FT /organism='Cytomegalovirus'

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10358"
BASE COUNT      126 a 131 c 123 g 132 t
ORIGIN
Query Match      25.4%; Score 127; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACATTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGTGATCGCG 76
Db 257 GACATTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGTGATCGCG 316
Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 317 TTTTGGCAGTACATCAATGGCGGTGGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 376
Qy 137 CACCCCA 143
Db 377 CACCCCA 383

RESULT 7
E59416
LOCUS
DEFINITION      Signal peptide.
ACCESSION      E59416
VERSION        E59416.1 GI:18622549
KEYWORDS       JP 2000354490-A/3.
SOURCE         unidentified
ORGANISM       unclassified.
REFERENCE      1 (bases 1 to 590)
AUTHORS       Hawkins,R. and Nakamura,M.
TITLE         Signal peptide
JOURNAL       Patent: JP 2000354490-A 3 26-DEC-2000;
COMMENT       OS cauliflower mosaic virus promoter
PN JP 2000354490-A/3
PD 26-DEC-2000
PF 15-JUN-1999 JP 1999168271
PR RICHARD HAWKINS,MICHI NAKAMURA
PI C12N15/09,C07K14/61,C12N1/13,C12P21/02,C12N15/00 CC
PC Key Location/Qualifiers
FH source 1..590
FT source /organism="cauliflower mosaic virus promoter".

FEATURES
source
Location/Qualifiers
1..590
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT      149 a 143 c 141 g 157 t
ORIGIN

Query Match      25.4%; Score 127; DB 6; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACATTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGTGATCGCG 76
Db 314 GACATTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGTGATCGCG 373
Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 374 TTTTGGCAGTACATCAATGGCGGTGGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 433
Qy 137 CACCCCA 143
Db 434 CACCCCA 440

RESULT 8
E51990
LOCUS
DEFINITION      Promoter, recombinant containing the same and utilization thereof.
ACCESSION      E51990
VERSION        E51990.1 GI:18629551
KEYWORDS       JP 2001000188-A/5.
SOURCE         Human herpesvirus 5
ORGANISM       Human herpesvirus 5
REFERENCE      1 (bases 1 to 599)
AUTHORS       Kubomura,M. and Saito,S.
TITLE         Promoter, recombinant containing the same and utilization thereof
JOURNAL       Patent: JP 2001000188-A 5 09-JAN-2001;
COMMENT       NIPPON ZEON CORP
OS Cytomegalovirus
PN JP 2001000188-A/5
PD 09-JAN-2001
PF 22-JUN-1999 JP 1999174804
PR MAYUMI KUBOMURA,SHUJI SAITO
PI C12N15/09,A61K31/00,A61K31/00,A61K39/02,A61K39/17,A61K39/215,
PC C07K14/125,
PC C07K14/165,C07K14/30,C12N7/00,C12N15/00
CC
CC
FH Key Location/Qualifiers
FT source 1..599
FT /organism="Cytomegalovirus".

FEATURES
source
Location/Qualifiers
1..599
/organism="Human herpesvirus 5"
/mol_type="genomic DNA"
/db_xref="taxon:10359"
BASE COUNT      152 a 144 c 144 g 159 t
ORIGIN

Query Match      25.4%; Score 127; DB 6; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACATTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGTGATCGCG 76
Db 319 GACATTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGTGATCGCG 378
Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 379 TTTTGGCAGTACATCAATGGCGGTGGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 438
Qy 137 CACCCCA 143
Db 439 CACCCCA 445

RESULT 9
AX360741
LOCUS
DEFINITION      Sequence 425 from Patent WO0202623.
ACCESSION      AX360741
VERSION        AX360741.1 GI:18676241
KEYWORDS
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      1
AUTHORS       Wang,T., McNeill,P.D., Wantanabe,Y., Carter,D., Henderson,R.A. and
Kalos,M.D.
TITLE         Compositions and methods for the therapy and diagnosis of lung
JOURNAL       Patent: WO 0202623-A 425 10-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..600

```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 153 a 152 c 141 g 154 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 76
Db 278 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 337
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
Db 338 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 397
QY 137 CACCCCA 143
Db 398 CACCCCA 404
RESULT 10
LOCUS AX256412 619 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 60 from Patent WO0170816.
ACCESSION AX256412
VERSION AX256412.1 GI:16075237
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Palli,S.R., Kapitskaya,M.Z. and Cress,D.E.
TITLE Ecdysone receptor-based inducible gene expression system
JOURNAL Patent: WO 0170816-A 60 27-SEP-2001;
ROHM AND HAAS COMPANY (US)
FEATURES
source
1..619
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 150 a 164 c 155 g 150 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 76
Db 254 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 313
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
Db 314 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 373
QY 137 CACCCCA 143
Db 374 CACCCCA 380
RESULT 11
LOCUS AX591955 738 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 24 from Patent WO0236760.
ACCESSION AX591955
VERSION AX591955.1 GI:27950185
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Lin,J., Yaver,D., Foster,D. and Holly,R.
TITLE Methods for expressing endogenous genes by restriction enzyme mediated integration
JOURNAL Patent: WO 0236760-A 24 10-MAY-2002;
Novozymes Biotech, Inc. (US); ZymoGenetics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..738
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 188 a 178 c 168 g 204 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 76
Db 335 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 394
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
Db 395 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 454
QY 137 CACCCCA 143
Db 455 CACCCCA 461
RESULT 12
LOCUS BD000145 777 bp DNA linear PAT 31-JAN-2002
DEFINITION Vector having stabilized sequence and eucaryotic host cell.
ACCESSION BD000145
VERSION BD000145.1 GI:18623224
KEYWORDS JP 2000308497-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 777)
AUTHORS Goman,C.M.
TITLE Vector having stabilized sequence and eucaryotic host cell
JOURNAL Patent: JP 2000308497-A 3 07-NOV-2000;
GENETIC INC
COMMENT
CS Unidentified
PN JP 2000308497-A/3
PD 07-NOV-2000
PF 17-APR-2000 JP 2000115248
PR 12-SEP-1986 US 907185,09-JUL-1987 US 071674 PI
CORNELIA MAXIN GOMAN
PC C12N15/09,C12N5/10,C12N15/00,C12N5/00
CC CC
FH Key Location/Qualifiers
FT source 1..773
/organism="Unidentified".
Location/Qualifiers
1..777
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 188 a 205 c 190 g 194 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 76
Db 343 GACATTTCTTCTGTCAGTACATCTAGCTATTAGTCTGCTATTACCATGGTGTGATCGGG 402

```

```
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 462
QY 137 CACCCCA 143
Db 463 CACCCCA 469

RESULT 13
I05430
LOCUS I05430 777 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 16 from Patent EP 0260148.
ACCESSION I05430
VERSION I05430.1 GI:591075
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 777)
AUTHORS Gorman,C.M.
TITLE Improved recombinant expression method, vector and transformed cells
JOURNAL Patent: EP 0260148-A2 16 16-MAR-1988;
FEATURES
source Location/Qualifiers
BASE COUNT 188 a 205 c 190 g 194 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 343 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 402
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 462
QY 137 CACCCCA 143
Db 463 CACCCCA 469

RESULT 14
I08105
LOCUS I08105 777 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0309237.
ACCESSION I08105
VERSION I08105.1 GI:589184
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 777)
AUTHORS Gorman,C.M.
TITLE A transient expression system for producing recombinant protein
JOURNAL Patent: EP 0309237-A1 3 29-MAR-1989;
FEATURES
source Location/Qualifiers
BASE COUNT 188 a 205 c 190 g 194 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
```

```
Db 343 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 402
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 462
QY 137 CACCCCA 143
Db 463 CACCCCA 469

RESULT 15
HSPPIELAB
LOCUS Human cytomegalovirus IE-1 promoter region. 875 bp DNA linear VRL 02-AUG-1993
DEFINITION Human cytomegalovirus IE-1 promoter region.
ACCESSION M64941
VERSION M64941.1 GI:330635
KEYWORDS enhancer region; promoter region.
SOURCE Human herpesvirus 5
ORGANISM Human herpesvirus 5
REFERENCE
1 (bases 1 to 875)
AUTHORS Lehner,R., Stammen,T. and Mach,M.
TITLE Comparative sequence analysis of human cytomegalovirus strains
JOURNAL J. Clin. Microbiol. 29 (11), 2494-2502 (1991)
MEDLINE 92129518
PUBMED 1663509
COMMENT Original source text: Human cytomegalovirus DNA.
Individual isolate: patient G.A.
major immediate early gene enhancer.
FEATURES
source Location/Qualifiers
1..875
/organism="Human herpesvirus 5"
/mol_type="genomic DNA"
/db_xref="taxon:10959"
/cell_line="infected human foreskin fibroblast"
BASE COUNT 227 a 209 c 201 g 238 t
ORIGIN Hind III F fragment of HCMV strain AD169.
Query Match 25.4%; Score 127; DB 14; Length 875;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 503 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 562
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
Db 563 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 622
QY 137 CACCCCA 143
Db 623 CACCCCA 629

Search completed: February 16, 2004, 01:25:45
Job time : 1992.02 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:07:11 ; Search time 161.871 Seconds
(without alignments)
8338.264 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500
Perfect score: 500
Sequence: 1 aagttgacattatgcact.....cactgggttgatgcgcgc 500

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Genesec 19Jun03.*
1: /SID1/gcgdata/genesec/genesecq-emb1/NA1980.DAT.*
2: /SID1/gcgdata/genesec/genesecq-emb1/NA1981.DAT.*
3: /SID1/gcgdata/genesec/genesecq-emb1/NA1982.DAT.*
4: /SID1/gcgdata/genesec/genesecq-emb1/NA1983.DAT.*
5: /SID1/gcgdata/genesec/genesecq-emb1/NA1984.DAT.*
6: /SID1/gcgdata/genesec/genesecq-emb1/NA1985.DAT.*
7: /SID1/gcgdata/genesec/genesecq-emb1/NA1986.DAT.*
8: /SID1/gcgdata/genesec/genesecq-emb1/NA1987.DAT.*
9: /SID1/gcgdata/genesec/genesecq-emb1/NA1988.DAT.*
10: /SID1/gcgdata/genesec/genesecq-emb1/NA1989.DAT.*
11: /SID1/gcgdata/genesec/genesecq-emb1/NA1990.DAT.*
12: /SID1/gcgdata/genesec/genesecq-emb1/NA1991.DAT.*
13: /SID1/gcgdata/genesec/genesecq-emb1/NA1992.DAT.*
14: /SID1/gcgdata/genesec/genesecq-emb1/NA1993.DAT.*
15: /SID1/gcgdata/genesec/genesecq-emb1/NA1994.DAT.*
16: /SID1/gcgdata/genesec/genesecq-emb1/NA1995.DAT.*
17: /SID1/gcgdata/genesec/genesecq-emb1/NA1996.DAT.*
18: /SID1/gcgdata/genesec/genesecq-emb1/NA1997.DAT.*
19: /SID1/gcgdata/genesec/genesecq-emb1/NA1998.DAT.*
20: /SID1/gcgdata/genesec/genesecq-emb1/NA1999.DAT.*
21: /SID1/gcgdata/genesec/genesecq-emb1/NA2000.DAT.*
22: /SID1/gcgdata/genesec/genesecq-emb1/NA2001A.DAT.*
23: /SID1/gcgdata/genesec/genesecq-emb1/NA2001B.DAT.*
24: /SID1/gcgdata/genesec/genesecq-emb1/NA2002.DAT.*
25: /SID1/gcgdata/genesec/genesecq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	5283	24	ABK10062 Expression vector
2	127	25.4	281	20	AAZ80727 Cytomegalovirus-de
3	127	25.4	282	20	AAZ80726 Cytomegalovirus-de
4	127	25.4	481	25	ABZ17986 S2 subtraction lib
5	127	25.4	500	19	AAV14005 Cytomegalovirus (C
6	127	25.4	589	20	AAV15163 Human cytomegalovi
7	127	25.4	590	22	AAH24425 Cauliflower mosaic
8	127	25.4	599	22	AAF84456 Cytomegalovirus (C

9	127	25.4	600	24	ABK16313 Human lung tumour
10	127	25.4	619	22	AAZ15664 Cauliflower mosaic
11	127	25.4	738	24	AAZ44414 CMV promoter / mou
12	127	25.4	763	21	AAZ95643 Cytomegalovirus de
13	127	25.4	777	9	AAZ82445 Partial expression
14	127	25.4	777	10	AAZ90697 DNA sequence of po
15	127	25.4	884	25	AAZ53783 CMV enhancer/Myelo
16	127	25.4	919	9	AAZ81527 Partial expression
17	127	25.4	919	10	AAZ90695 DNA sequence of po
18	127	25.4	930	7	AAZ60156 Enhancer for eukar
19	127	25.4	930	14	AAZ53550 HCMV (AD169) major
20	127	25.4	930	14	AAZ43524 Sequence of HCMV(A
21	127	25.4	930	18	AAZ77193 HCMV immediate ear
22	127	25.4	987	24	AAZ37261 Adeno-associated v
23	127	25.4	996	20	AAZ25629 Human sonic hedgeh
24	127	25.4	996	20	AAZ25113 Human sonic hedgeh
25	127	25.4	996	20	AAZ16193 Human Shh gene PCR
26	127	25.4	996	21	AAZ27888 Human sonic hedgeh
27	127	25.4	996	21	AAZ30286 Human sonic hedgeh
28	127	25.4	996	22	AAZ166796 Human sonic hedgeh
29	127	25.4	996	22	AAZ28445 Nucleotide sequenc
30	127	25.4	996	24	AAZ23811 Human sonic hedgeh
31	127	25.4	1070	19	AAZ58058 Plasmid CMV-delRT
32	127	25.4	1104	19	AAZ03231 DNA encoding CTLA-
33	127	25.4	1215	17	AAZ23597 CMV-PC-PSA promote
34	127	25.4	1216	17	AAZ29595 CMV-PSA promoter.
35	127	25.4	1467	23	ABX04557 DNA encoding Synec
36	127	25.4	1632	24	AAZ44408 CMV promoter / mou
37	127	25.4	1715	24	AAZ38382 Chimeric sequence
38	127	25.4	1767	24	AAZ38380 Wild-type human AD
39	127	25.4	1767	24	AAZ38391 Human ovarian anti
40	127	25.4	1787	24	ABQ55020 Transgenic APP DNA
41	127	25.4	1813	21	AAZ89475 Reporter gene cons
42	127	25.4	1814	21	AAA47118 DNA encoding N-ter
43	127	25.4	2043	23	ABX04567 Promoter-enhancer
44	127	25.4	2133	10	AAZ91042 CMVintA promoter a
45	127	25.4	2196	15	AAZ74211

ALIGNMENTS

RESULT 1

ABK10062
ID ABK10062 standard; DNA; 5283 BP.

XX ABK10062;

AC 21-MAY-2002 (first entry)

XX Expression vector construct pVGI.1 containing VEGF-2 insert.

DE Human; vascular endothelial growth factor 2; VEGF-2; ds; pVGI.1;
KW Chronic limb ischaemia; myocardial ischaemia; autoimmune disorder; pHE4a;
KW allergic reaction; organ rejection; inflammatory condition; arrhythmia;
KW hyperproliferative disorder; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cardiovascular disorder; embolism;
KW heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.
XX Synthetic.

XX WO200211769-A1.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US24658.

XX 04-AUG-2000; 2000US-223276P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Coleman TA;

XX

DR WPI; 2002-217153/27.
XX Isolated nucleic acid having expression vector construct with vascular
PT endothelial growth factor-2 insert, useful for treating chronic limb
PT ischaemia or myocardial ischaemia, autoimmune disorders and allergic
PT conditions -
XX
XX Disclosure; Fig 31; 241pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising pVGI.1
CC expression vector construct containing the vascular endothelial growth
CC factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host
CC cell by transducing, transforming or transfecting a host cell with the
CC DNA and for treating a patient having chronic limb ischaemia or
CC myocardial ischaemia, or a disease or disorder selected from autoimmune
CC disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or
CC conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.
CC Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),
CC diseases due to viral, bacterial, fungal or parasitic infection,
CC cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve
CC diseases, aneurysms, arterial occlusive disorders and embolism. This
CC sequence represents the pVGI.1 expression vector containing the VEGF-2
CC insert.
XX
XX Sequence 5283 BP; 1244 A; 1422 C; 1417 G; 1200 T; 0 other;

Query Match 100.0%; Score 500; DB 24; Length 5283;
Best Local Similarity 100.0%; Pred. No. 4.1e-262;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCTTGACCTTATGCGACTTCTCTACTTGGCAGTACATCTAGCTATTAGTCTATCCTAT 60
Db 1 AAGCTTGACCTTATGCGACTTCTCTACTTGGCAGTACATCTAGCTATTAGTCTATCCTAT 60
Qy 61 TACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGTTGACTCAG 120
Db 61 TACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGTTTGGCTCAG 120
Qy 121 GGGATTTCCAGTCTCCACCCCACTGACGTCATGGAGTTCTTTTGGCACCATAATCA 180
Db 121 GGGATTTCCAGTCTCCACCCCACTGACGTCATGGAGTTCTTTTGGCACCATAATCA 180
Qy 181 ACAGACTTTCCAAATGTGCTAACAACTCCGCCCACTTACGCAAAATGGCGGTAGGGC 240
Db 181 ACAGACTTTCCAAATGTGCTAACAACTCCGCCCACTTACGCAAAATGGCGGTAGGGC 240
Qy 241 AACATGCTTATGTAACGGTGAGTTAGCAATGCTTATAGGAGAGAAACACCGTG 300
Db 241 AACATGCTTATGTAACGGTGAGTTAGCAATGCTTATAGGAGAGAAACACCGTG 300
Qy 301 CATGCCGATTGGTGGAGTAAGGTGGTATGATCGTGTATGATCGTGCCTTGTAGGAAG 360
Db 301 CATGCCGATTGGTGGAGTAAGGTGGTATGATCGTGTATGATCGTGCCTTGTAGGAAG 360
Qy 361 GCAACAGACGGGTCTAACAGGATGGACGAAACACTGAAATTCGGCATTGACAGATATT 420
Db 361 GCAACAGACGGGTCTAACAGGATGGACGAAACACTGAAATTCGGCATTGACAGATATT 420
Qy 421 GTATTTAAGTCCCGAGCTCGATACATAAAGCGCATTTGACCATTCACCACTTGGTGTG 480
Db 421 GTATTTAAGTCCCGAGCTCGATACATAAAGCGCATTTGACCATTCACCACTTGGTGTG 480
Qy 481 CACCTGGTGGATCGATC 500
Db 481 CACCTGGTGGATCGATC 500

RESULT 2
AAx80727
ID AAx80727 standard; DNA; 281 BP.
XX
AC AAx80727;
XX

DT 19-OCT-1999 (first entry)
XX Cytomegalovirus-derived promoter element-2 (-299 to -19).
DE
XX
XX Cytomegalovirus-derived promoter element-2; CMV-derived promoter element;
KW Cytomegalovirus immediate early promoter; human alpha-galactosidase gene;
KW persistent transgene expression; expression cassette; human CFTR gene;
KW human cystic fibrosis transmembrane conductance regulator gene;
KW adenoviral vector; E4 region; transgene; plasmid; ds.
XX
OS Cytomegalovirus.
XX
XX Key Location/Qualifiers
FH promoter 1..281
FT /*tag= a
FT /label= CMV_derived_promoter_element
XX
PN W09936557-A1.
XX
XX 22-JUL-1999.
PD
XX 15-JAN-1999; 99WO-US00915.
PF
XX 16-JAN-1998; 98US-0071673.
PR
XX (GENZ) GENZYME CORP.
PA
XX Armentano D, Marshall J, Yew N;
XX WPI; 1999-478911/40.
XX
XX A cytomegalovirus (CMV)-derived promoter element and human albumin
PT gene enhancer elements useful for persistent gene expression
XX
XX Claim 3; Page 49; 56pp; English.
XX
XX The present sequence is a promoter element derived from the
CC cytomegalovirus (CMV) immediate early promoter. The promoter element
CC is used to achieve persistent expression of an operably linked transgene,
CC especially human cystic fibrosis transmembrane conductance regulator
CC gene (CFTR gene) or human alpha-galactosidase gene, in a target cell. The
CC expression cassette comprising the promoter element operably linked to a
CC transgene is delivered to a target cell via a plasmid or an adenoviral
CC vector lacking the E4 region. As the promoter element functions
CC independently of the viral E4 region, this allows the use of vectors
CC containing reduced viral genomes thereby increasing the carrying capacity
CC of the vector while decreasing the potential for host immune reaction.
XX
XX Sequence 281 BP; 67 A; 64 C; 73 G; 77 T; 0 other;
SQ
Query Match 25.4%; Score 127; DB 20; Length 281;
Best Local Similarity 100.0%; Pred. No. 8.1e-59;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCTATTCATCGTATTACATGGTATCGGG 76
Db 31 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCTATTCATCGTATTACATGGTATCGGG 90
Qy 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
Db 91 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 150
Qy 137 CACCCCA 143
Db 151 CACCCCA 157
XX
XX RESULT 3
XX AAx80726
ID AAx80726 standard; DNA; 282 BP.
XX
AC AAx80726;
XX

```

DT 19-OCT-1999 (first entry)
XX
XX DE Cytomegalovirus-derived promoter element-1 (-295 to -14).
XX
XX KW Cytomegalovirus-derived promoter element-1; CMV-derived promoter element;
XX KW Cytomegalovirus immediate early promoter; human alpha-galactosidase gene;
XX KW persistent transgene expression; expression cassette; human CFTR gene;
XX KW human cystic fibrosis transmembrane conductance regulator gene;
XX KW adenoviral vector; E4 region; transgene; plasmid; ds.
XX
XX OS Cytomegalovirus.
XX
XX FH Key
XX FH Location/Qualifiers
XX FT Promoter
XX FT 1..282
XX FT /tag= a
XX FT /label= CMV_derived_promoter_element
XX
XX PN WO9936557-A1.
XX
XX PD 22-JUL-1999.
XX
XX PF 15-JAN-1999; 99WO-US00915.
XX
XX PR 16-JAN-1998; 98US-0071673.
XX
XX PA (GENZ ) GENZYME CORP.
XX
XX PI Armentano D, Marshall J, Yew N;
XX
XX WPI; 1999-478911/40.
XX
XX A cytomagalovirus (CMV)-derived promoter element and human albumin
XX gene enhancer elements useful for persistent gene expression
XX
XX Claim 3; Pages 48-49; 56pp; English.
XX
XX The present sequence is a promoter element derived from the
XX cytomagalovirus (CMV) immediate early promoter. The promoter element
XX is used to achieve persistent expression of an operably linked transgene,
XX especially human cystic fibrosis transmembrane conductance regulator
XX gene (CFTR gene) or human alpha-galactosidase gene, in a target cell. The
XX expression cassette comprising the promoter element operably linked to a
XX transgene is delivered to a target cell via a plasmid or an adenoviral
XX vector lacking the E4 region. As the promoter element functions
XX independently of the viral E4 region, this allows the use of vectors
XX containing reduced viral genomes thereby increasing the carrying capacity
XX of the vector while decreasing the potential for host immune reaction.
XX
XX SQ Sequence 282 BP; 68 A; 65 C; 72 G; 77 T; 0 other;
XX
XX Query Match 25.4%; Score 127; DB 20; Length 282;
XX Best Local Similarity 100.0%; Pred. No. 8.1e-59;
XX Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 76
XX DB 27 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 86
XX
XX QY 77 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
XX DB 87 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 146
XX
XX QY 137 CACCCCA 143
XX DB 147 CACCCCA 153
XX
XX RESULT 4
XX ABZ17986
XX ID ABZ17986 standard; cDNA; 481 BP.
XX
XX AC ABZ17986;
XX
XX DT 18-JUN-1998 (first entry)
XX
XX DE S2 subtraction library cancer related clone SEQ ID NO:412.
XX
XX KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
XX KW immune response; virology; immunology; microbiology; molecular biology;
XX KW recombinant DNA technology; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200278516-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US10421.
XX
XX PR 30-MAR-2001; 2001US-280255P.
XX PR 28-AUG-2001; 2001US-315563P.
XX PR 09-JAN-2002; 2002US-347313P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
XX WPI; 2003-058387/05.
XX
XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
XX preventing and treating cancer expressing CT or CP mRNA antigens, and
XX in virology, immunology, microbiology, molecular biology and
XX recombinant DNA techniques -
XX
XX Claim 1; SEQ ID 412; 207pp; English.
XX
XX ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
XX ABP5446 to ABP5447 represent protein (II) sequences, from the present
XX invention. (I) and (II) have cytostatic activity and can be used in gene
XX therapy and vaccines. (I), (II), antibodies and compositions from the
XX present invention are useful for diagnosing, preventing and treating
XX cancer, which expresses CT or CP mRNA antigens. They are useful for
XX stimulating immune response. They can also be useful in virology,
XX immunology, microbiology, molecular biology and recombinant DNA
XX techniques.
XX
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 481 BP; 121 A; 119 C; 109 G; 132 T; 0 other;
XX
XX Query Match 25.4%; Score 127; DB 25; Length 481;
XX Best Local Similarity 100.0%; Pred. No. 8e-59;
XX Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 76
XX DB 315 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 374
XX
XX QY 77 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 136
XX DB 375 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTC 434
XX
XX QY 137 CACCCCA 143
XX DB 435 CACCCCA 441
XX
XX RESULT 5
XX AAV14005
XX ID AAV14005 standard; cDNA; 500 BP.
XX
XX AC AAV14005;
XX
XX DT 18-JUN-1998 (first entry)
XX

```

DE Cytomegalovirus (CMV)-MIE wild-type promoter sequence.
XX Gfi-1; transcription repressor; mutation; cytokine; gene therapy;
KW growth factor independence-1; cytomegalovirus; promoter; ss.
XX
OS Human cytomegalovirus.
XX Key Location/Qualifiers
FH misc_binding 331..342
FT /tag= a
FT /note= "putative Gfi-1 binding site"
FT misc_feature 337
FT /tag= b
FT /note= "G can be substituted by A"
FT misc_feature 338
FT /tag= c
FT /note= "A can be substituted by C"
FT misc_feature 339
FT /tag= d
FT /note= "T can be substituted by G"
FT misc_binding 388..399
FT /tag= e
FT /note= "putative Gfi-1 binding site"
FT misc_feature 391
FT /tag= f
FT /note= "A can be substituted by C"
FT misc_feature 392
FT /tag= g
FT /note= "T can be substituted by C"
FT misc_feature 393
FT /tag= h
FT /note= "C can be substituted by T"
FT CAAT_signal 438..442
FT /tag= i
FT TATA_signal 471..476
FT /tag= j
XX
PN WO9748720-A1.
XX
XX 24-DEC-1997.
PD
XX 17-JUN-1997; 97WO-US10486.
XX
XX 17-JUN-1996; 96US-0019808.
XX
XX (FOX-) FOX CHASE CANCER CENT.
XX
XX Grimes HL, Tsichlis P, Zweidler-Mckay P;
XX WPI; 1998-063073/06.
XX
XX DNA containing inactive, mutated binding site for Gfi-1
XX transcription repressor - used to increase gene expression in vitro
XX or in vivo, e.g. in gene therapy
XX
XX Disclosure; Fig 3A; 44pp; English.
XX
XX This is the wild-type cytomegalovirus (CMV)-MIE promoter sequence
XX comprising two putative Gfi-1 (growth factor independence-1)
XX transcription repressor binding sites. This promoter can be mutated at
XX the Gfi-1 binding sites to be used in a novel isolated DNA construct
XX which contains at least one mutated binding site for a Gfi-1
XX transcription repressor that hinders or prevents binding of Gfi-1 to
XX this site. The expression vector contains an expression regulatory
XX segment that contains at least one copy of the sequences shown in
XX AAV19671 to AAV19685 linked operably to a coding segment selected from a
XX group of cytokines, interleukins, interferons, growth factors and proto-
XX oncogenes, and an isolated DNA molecule containing one of two 500 base
XX pair sequences shown in AAV14003 and AAV14004. Altering the binding site
XX increases expression of these genes controlled by regulators that
XX include binding sites for Gfi-1 both in cultured cells and in vivo (for
XX gene therapy or DNA vaccination). A vector containing a normal gene
XX under control of a regulator with the mutated binding site can be

CC administered to a patient having a disease associated with an aberrant
XX form of the gene.
XX
SQ Sequence 500 BP; 121 A; 126 C; 125 G; 128 T; 0 other;
Query Match 25.4%; Score 127; DB 19; Length 500;
Best Local Similarity 100.0%; Pred. No. 8e-59;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGTTGATGCGG 76
DB 231 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGTTGATGCGG 290
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTACGGGGATTTCCAAGTCTC 136
DB 291 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTACGGGGATTTCCAAGTCTC 350
QY 137 CACCCCA 143
DB 351 CACCCCA 357
RESULT 6
AA15163/C
ID AAX15163 standard; DNA; 589 BP.
XX
AC AAX15163;
XX
DT 22-APR-1999 (first entry)
XX
DE Human cytomegalovirus promoter region.
XX
KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis;
KW Human cytomegalovirus promoter region; ss.
XX
OS Human cytomegalovirus.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
XX 10-JUN-1998; 98WO-US11927.
XX
XX 14-MAY-1998; 98US-0079030.
XX
XX 13-JUN-1997; 97US-0874807.
XX
XX (BAYU) BAYLOR COLLEGE MEDICINE.
XX
XX Guevara JG, Hoogveen RC, Moore JP;
XX WPI; 1999-070331/06.
XX
XX Composition comprising nucleic acid bound to LDL or VLDL lipoprotein
XX - used for delivering nucleic acid to cells for gene therapy and
XX antisense treatment
XX
XX Example 4; Page 74; 293pp; English.
XX
XX The present sequence represents the Human cytomegalovirus promoter
XX region. The compositions of the invention bind to this sequence.
XX The specification describes a composition that comprises low
XX density lipoprotein (LDL) and apolipoproteins for the binding and
XX in vivo transport of nucleic acids. The composition is used to deliver
XX nucleic acids to eukaryotic cells, in vivo or in vitro, for expressing
XX a therapeutic polypeptide or antisense molecule (or ribozyme).
XX Specifically they are used for gene therapy of cancers (particularly
XX non-small cell lung carcinoma), diabetes, cystic fibrosis and
XX arteriosclerosis.
XX
SQ Sequence 589 BP; 157 A; 140 C; 143 G; 149 T; 0 other;

Query Match 25.4%; Score 127; DB 20; Length 589;
 Best Local Similarity 100.0%; Pred. No. 8e-59; Mismatches 0; Indels 0; Gaps 0;
 Matches 127; Conservative 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
 DB 276 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 217

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 136
 DB 216 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 157

QY 137 CACCCCA 143
 DB 156 CACCCCA 150

RESULT 7
 AAH24425
 ID AAH24425 standard; DNA; 590 BP.
 XX
 AC AAH24425;
 XX
 DT 02-AUG-2001 (first entry)
 XX
 DE Cauliflower mosaic virus promoter.
 XX
 DE Cauliflower mosaic virus; promoter; signal peptide; chloroella;
 KW gene expression; protein production; human growth hormone; GS.
 KW
 XX Cauliflower mosaic virus.
 OS
 XX JP2000354490-A.
 PN
 XX 26-DEC-2000.
 PD
 XX 15-JUN-1999; 99JP-0168271.
 PF
 XX 15-JUN-1999; 99JP-0168271.
 PR
 XX (TOYT) TOYOTA JIDOSHA KK.
 PA
 XX WPI; 2001-275809/29.
 DR
 XX
 XX New signal peptides useful for the preparation of human growth hormone
 PT and transformed chloroella -
 PT
 XX
 XX Example 1; Page 12; 15pp; Japanese.
 PS
 XX The present sequence is provided in a specification relating to signal
 CC peptides for expression and secretion of a protein in chloroella. The
 CC peptides are of the formula:
 CC Met-Ala-Asn-Lys-X₁-1-(Leu)_n-X₂-Ala-Ser-Gly.
 CC X₁ = Ser or Leu;
 CC n = an integer of 5-15;
 CC X₂ = Gly-Ser-Leu or Pro-Leu-Ala.
 CC The signal peptides are useful in the preparation of human growth
 CC hormone and transformed chloroella. Signal peptides, DNA encoding the
 CC peptides, gene expression cassettes, recombinant vectors containing the
 CC cassettes, and transformants having the vectors are provided. The
 CC present sequence is a promoter which may be used in the invention.
 XX
 XX Sequence 590 BP; 149 A; 143 C; 141 G; 157 T; 0 other;
 SQ

Query Match 25.4%; Score 127; DB 22; Length 590;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
 DB 314 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 373

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 136

Db 374 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 433
 QY 137 CACCCCA 143
 DB 434 CACCCCA 440

RESULT 8
 AAF84456
 ID AAF84456 standard; DNA; 599 BP.
 XX
 AC AAF84456;
 XX
 DT 25-JUN-2001 (first entry)
 XX
 DE Cytomegalovirus (CMV) immediate-early (IE) promoter enhancer.
 XX
 KW Cytomegalovirus; CMV; immediate-early; IE promoter enhancer;
 KW Pec promoter; chicken beta-actin promoter;
 KW recombinant vector; genetic vaccine; gene therapy; ds.
 XX
 OS Human cytomegalovirus.
 XX
 PN JP2001000188-A.
 XX
 PD 09-JAN-2001.
 XX
 PF 22-JUN-1999; 99JP-0174804.
 PR 22-JUN-1999; 99JP-0174804.
 XX
 XX (JAPG) NIPPON ZEON KK.
 PA
 XX WPI; 2001-285232/30.
 DR
 XX
 XX New DNA molecule for use as a promoter for preparing a recombinant
 PT containing the DNA which is used for preparing a vaccine -
 PT
 XX
 PS Example 1; Page 12; 15pp; Japanese.
 XX
 CC The invention relates to a 5' and a 3' fragment of the chicken beta-actin
 CC promoter (AAF84452 and AAF84453, respectively), which exhibit promoter
 CC activity. The chicken beta-actin promoter fragments may be used to drive
 CC expression of a heterologous gene in a recombinant vector used as the
 CC active component of a genetic vaccine. The present sequence represents
 CC a cytomegalovirus (CMV) immediate-early (IE) promoter enhancer, a
 CC portion of which was used in the construction of the chimeric Pec
 CC promoter (AAF84454) in an exemplification of the invention.
 XX
 XX Sequence 599 BP; 152 A; 144 C; 144 G; 159 T; 0 other;
 SQ

Query Match 25.4%; Score 127; DB 22; Length 599;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
 DB 319 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 378

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 136
 DB 379 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 438

QY 137 CACCCCA 143
 DB 439 CACCCCA 445

RESULT 9
 ABK16313
 ID ABK16313 standard; cDNA; 600 BP.
 XX

AC ABK16313;
 XX 14-MAR-2002 (first entry)
 DT
 DE Human lung tumour protein encoding cDNA #425.
 XX
 XX Human lung tumour protein; lung cancer; immunostimulant; cytostatic; ss;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 XX WO200202623-A2.
 PN
 XX 10-JAN-2002.
 PD
 XX
 XX 28-JUN-2001; 2001WO-US20975.
 PF
 XX 29-JUN-2000; 2000US-215696P.
 PR
 XX 22-AUG-2000; 2000US-227142P.
 PR
 XX 06-SEP-2000; 2000US-230481P.
 PR
 XX 21-DEC-2000; 2000US-257729P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Wang T, McNeill PD, Wantanabe Y, Carter D, Henderson RA, Kalos MD;
 PI
 XX WPI; 2002-106602/14.
 DR
 XX
 XX Polynucleotides encoding lung tumour polypeptides, useful for treating
 PT lung cancer or stimulating an immune response -
 XX
 XX Claim 1; Page 261-262; 277pp; English.
 XX
 XX The invention relates to polynucleotides encoding lung tumour
 CC polypeptides. Compositions comprising the lung tumour polypeptides,
 CC polynucleotides, fusion proteins, antibodies to the polypeptides, T cell
 CC populations, or antigen presenting cells that express the lung tumour
 CC polypeptides, are useful for treating lung cancer or stimulating an
 CC immune response. The polynucleotides and polypeptides are also useful in
 CC the diagnosis and monitoring of lung cancer. Sequences ABK16343
 CC represent cDNA molecules encoding the polypeptides of the invention and
 CC PCR primers used to amplify the cDNA.
 XX
 XX Sequence 600 BP; 153 A; 152 C; 141 G; 154 T; 0 other;
 SQ
 Query Match 25.4%; Score 127; DB 24; Length 600;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCTATCGCTATTACCATGGTGATCGG 76
 DB 278 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCTATCGCTATTACCATGGTGATCGG 337
 QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
 DB 338 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 397
 QY 137 CACCCCA 143
 DB 398 CACCCCA 404
 RESULT 10
 AAS15664
 ID AAS15664 standard; DNA; 619 BP.
 XX
 XX AAS15664;
 AC
 XX 29-JAN-2002 (first entry)
 DT
 XX Cauliflower mosaic virus promoter.
 DE
 XX Ecdysone receptor; EcR; ligand binding domain; ds;
 XX
 KW

KW retinoid X receptor; RXRalpha; DNA-binding domain; CMV promoter;
 KW transactivation domain; nuclear receptor; ultraspicrable; gene therapy;
 KW protein production; antibody production; high throughput screening;
 KW HTS; transgenic plant; transgenic animal.
 XX
 OS Cauliflower mosaic virus.
 XX
 XX WO200170816-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX
 XX 21-MAR-2001; 2001WO-US09050.
 PF
 XX 22-MAR-2000; 2000US-191355P.
 PR
 XX 20-FEB-2001; 2001US-269799P.
 PR
 XX (ROHM) ROHM & HAAS CO.
 PA
 XX Palli SR, Kapitskaya MZ, Cress DE;
 PI
 XX WPI; 2001-656841/75.
 DR
 XX Ecdysone and retinoid X receptor based inducible gene expression
 PT systems for use in e.g. gene therapy, large scale production of
 PT proteins and cell-based high-throughput screening assays -
 XX
 XX Example 1; Page 140-141; 144pp; English.
 XX
 XX The invention relates to Ecdysone and retinoid X receptor based inducible
 CC gene expression systems useful for modulating gene expression in host
 CC cells. The gene expression system encodes a polypeptide with a
 CC DNA-binding domain recognizes a response element associated with a gene
 CC whose expression is to be modulated and/or a ligand binding domain (LBD)
 CC comprising a LBD from a nuclear receptor and a second gene expression
 CC cassette capable of being expressed in a host cell comprising a
 CC polynucleotide sequence encoding a second polypeptide comprising a trans-
 CC activation domain and/or a LBD comprising a LBD from a nuclear receptor
 CC other than ultraspicrable (USP) (the trans-activation domain is from a
 CC nuclear receptor other than an ecdysone receptor, a retinoid X receptor
 CC or a USP receptor and the LBDs from the first and second polypeptides are
 CC different and dimerise). The ecdysone and retinoid X receptor based
 CC inducible gene expression systems useful for modulating gene expression
 CC in host cells, for use in gene therapy, large scale production of
 CC proteins and antibodies, cell-based high-throughput screening
 CC assays (HTS), functional genomic and regulation of traits in transgenic
 CC plants and animals. The present sequence represents a cauliflower mosiac
 CC virus, CMV, promoter which is used in an inducible gene expression
 CC system of the invention.
 XX

SQ Sequence 619 BP; 150 A; 164 C; 155 G; 150 T; 0 other;
 Query Match 25.4%; Score 127; DB 22; Length 619;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCTATCGCTATTACCATGGTGATCGG 76
 DB 254 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCTATCGCTATTACCATGGTGATCGG 313
 QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
 DB 314 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 373
 QY 137 CACCCCA 143
 DB 374 CACCCCA 380

RESULT 11
 AAL44414
 ID AAL44414 standard; DNA; 738 BP.
 XX
 XX AAL44414;
 AC

```

XX 31-OCT-2002 (first entry)
XX CMV promoter / mouse G-CSF receptor (G-CSFR) gene 9 - chimeric sequenc.
XX Mouse; murine; ds; gene of interest isolation; CMV; gene; promoter;
XX novel gene discovery; novel gene isolation; large library production;
XX G-CSF receptor; G-CSFR; chimeric.
XX Chimeric - Mus sp.
XX Chimeric - Cytomegalovirus.
XX WO200236760-A2.
XX 10-MAY-2002.
XX 30-OCT-2001; 2001WO-US45045.
XX 30-OCT-2000; 2000US-0702177.
XX (NOVO) NOVOZYMES BIOTECH INC.
XX (ZYMO) ZYMOGENETICS.
XX Lin J, Yaver D, Foster D, Holly R;
XX WPI; 2002-566503/60.
XX Production of a mutant mammalian cell for isolating genes, comprising
XX integrating a nucleic acid construct into the cell genome at sites
XX generated by the restriction enzyme, and selecting cell with trait of
XX interest.
XX Example 11; Fig 14; 106pp; English.
XX The invention comprises a method for producing a mutant mammalian cell.
XX The method involves introducing a restriction enzyme and a nucleic acid
XX regulatory sequence into mammalian cells - for integrating the nucleic
XX acid construct into the mammalian cell genome at sites generated by the
XX restriction enzyme. Mutant mammalian cells having a trait of interest can
XX then be selected. The method of the invention is useful for isolating a
XX gene controlling a trait of interest from a mammalian cell. The method is
XX useful for discovering and isolating new genes. The method of the
XX invention can be used to create large libraries of mammalian cells which
XX have a low transfection efficiency. The method of the invention is also
XX suitable for over-expressing a known endogenous gene that is expressed
XX poorly. The present DNA sequence represents a chimeric gene sequence
XX comprising a Cytomegalovirus (CMV) promoter and a murine G-CSF receptor
XX (G-CSFR) gene.
XX Sequence 738 BP; 188 A; 178 C; 168 G; 204 T; 0 other;
XX
XX Query Match 25.4%; Score 127; DB 24; Length 738;
XX Best Local Similarity 100.0%; Pred. No. 8e-59;
XX Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGTCATGCGG 76
XX DB 335 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGTCATGCGG 394
XX
XX QY 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCAACGGGGATTTCCAAAGTCTC 136
XX DB 395 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCAACGGGGATTTCCAAAGTCTC 454
XX
XX QY 137 CACCCCA 143
XX DB 455 CACCCCA 461
XX
XX RESULT 12
XX AAZ95643
XX ID AAZ95643 standard; DNA; 763 BP.
XX AC AAZ95643;

```

```

XX 08-JUN-2000 (first entry)
XX Cytomegalovirus derived promoter sequence SEQ ID NO:1.
XX
XX Cytomegalovirus; CMV; promoter; human growth hormone; PGH; ovine;
XX prostaglandin G/H synthase; plasmid; prostanoind; circular;
XX cardiant; thrombolytic; antiinflammatory; gene therapy;
XX pulmonary embolus; myocardial infarction; lung disease;
XX adult respiratory distress syndrome; ds.
XX Human cytomegalovirus.
XX OS
XX US6030638-A.
XX 29-FEB-2000.
XX 02-JUN-1995; 95US-0459493.
XX 19-AUG-1991; 91US-0746941.
XX 21-JUN-1993; 93US-0080221.
XX (UYVA-) UNIV VANDERBILT.
XX Meyrick B, Canonico A, Brigham K, Conary JT;
XX WPI; 2000-205193/18.
XX Increasing prostanoind production in vivo comprising delivering and
XX hyperexpressing a prostaglandin synthase gene in cells, useful for
XX treating e.g. myocardial infarction and general lung disease -
XX Disclosure; Column 7-10; 11pp; English.
XX
XX A method has been developed of increasing prostanoind production in vivo
XX comprising delivering and hyperexpressing a prostaglandin synthase gene
XX in cells. The method can be used for gene therapy. The method is useful
XX for the treatment of diseases such as pulmonary embolus, myocardial
XX infarction and general lung disease (e.g. adult respiratory distress
XX syndrome). Generally the present invention provides a plasmid comprising
XX a cytomegalovirus (CMV) derived promoter sequence driving the coding
XX region for ovine prostaglandin G/H synthase (PGH). In order to increase
XX the expression of the gene, the construct also contains a short
XX translation augmenting sequence and a portion of the 3' untranslated
XX region from the human growth hormone (hGH) gene. The present sequence
XX represents the CMV promoter sequence used in the exemplification of the
XX present invention.
XX Sequence 763 BP; 203 A; 174 C; 164 G; 222 T; 0 other;
XX
XX Query Match 25.4%; Score 127; DB 21; Length 763;
XX Best Local Similarity 100.0%; Pred. No. 8e-59;
XX Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGTCATGCGG 76
XX DB 492 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGTCATGCGG 551
XX
XX QY 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCAACGGGGATTTCCAAAGTCTC 136
XX DB 552 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCAACGGGGATTTCCAAAGTCTC 611
XX
XX QY 137 CACCCCA 143
XX DB 612 CACCCCA 618
XX
XX RESULT 13
XX AA82445
XX ID AA82445 standard; DNA; 777 BP.
XX AC AA82445;

```


DT 06-MAR-1992 (first entry)
 XX Partial expression vector pF8C8SS sequence.
 DE Expression vector; promoter; heterologous protein; mammalian vector;
 KW ds.
 XX Cytomegalovirus.
 OS Key Location/Qualifiers
 FH misc_feature 2
 FT /tag= a
 FT /note= "from pPMLCMV beginning to HindIII,
 FT enhancers and promoter"
 FT 612
 FT /tag= b
 FT /note= "starts here"
 PN EP260148-A.
 XX 16-MAR-1988.
 PD 11-SEP-1987; 87EP-0308060.
 PF 09-JUL-1987; 87US-0071674.
 PR 12-SEP-1986; 86US-0907185.
 PR 25-SEP-1987; 87US-0101712.
 XX (GETH) GENENTECH INC.
 PA Gorman CM;
 XX WPI; 1988-072583/11.
 DR Continuous expression of heterologous protein in eukaryotic cells
 PT - by transfecting with vector contg. stabilising sequence,
 PT promoter, structural gene, poly:adenylate sequence and
 PT termination site
 XX Example; Fig 19; 48pp; English.
 PS In the expression vectors of the invention, the promoter is
 XX specifically from the immediate early gene of human cytomegalovirus
 CC (CMV) or from simian virus (SV40). The stabilizing SQ contains 1 or
 CC 2 splice donor-intron-acceptor SQs, esp. from the specified CMV gene,
 CC or the intron is from CMV or the Ig variable region, while the
 CC splice-acceptor component corresponds to the Ig acceptor SQ. The
 CC vector may also include (upstream of the promoter) an enhancer, esp.
 CC also from SV40.
 XX Sequence 777 BP; 188 A; 205 C; 190 G; 194 T; 0 other;
 SQ Query Match 25.4%; Score 127; DB 9; Length 777;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTTGGCAGTACATCTACCTATTAGTCATCGCTATTACATGGTATCGGG 76
 DB 343 GACTTTCCTACTTGGCAGTACATCTACCTATTAGTCATCGCTATTACATGGTATCGGG 402
 QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGAATCAACGGGATTTCCAGTCTC 136
 DB 403 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGAATCAACGGGATTTCCAGTCTC 462
 QY 137 CACCCCA 143
 DB 463 CACCCCA 469
 RESULT 14
 AAN90697
 ID AAN90697 standard; DNA; 777 BP.
 XX

AC AAN90697;
 XX 25-MAR-2003 (updated)
 DT 15-JUL-1990 (first entry)
 XX DNA sequence of portion of plasmid pF8C8SS containing the cytomegalovirus
 DE enhancer and promoter and an engineered stabilising.
 XX Plasmid pF8C8SS; cytomegalovirus; SV40; enhancer; promoter;
 KW stabilising sequence; splice donor intron sequence; Ig region;
 KW splice acceptor sequence; human embryonic kidney cells (293); JW2.
 XX Cytomegalovirus; SV40.
 OS Key Location/Qualifiers
 FH promoter 1..732
 FT /tag= a
 FT /note= "cytomegalovirus enhancer, promoter and leader"
 FT misc_feature 733..736
 FT /tag= b
 FT /note= "stabilising sequence including the engineered
 FT splice donor and acceptor sequence"
 FT 737..777
 FT /tag= c
 FT misc_RNA 612
 FT /tag= d
 FT /note= "Begin RNA"
 XX EP309237-A.
 PN 29-MAR-1989.
 XX 22-SEP-1988; 89EP-0308784.
 PR 25-SEP-1987; 87US-0101712.
 XX (GETH) GENENTECH INC.
 PA Gorman CM;
 XX WPI; 1989-095394/13.
 DR Transient expression system for recombinant proteins -
 CC comprising eukaryotic host cell transfected with vector
 CC encoding trans-activating protein and expression vector
 XX Disclosure; 44 pp.; English.
 PS Expression vector pF8C8SS contains the cytomegalovirus enhancer and
 CC promoter, an engineered stabilising sequence, the cDNA encoding factor
 CC VIII and the SV40 polyadenylation site. The entire intron region
 CC including donor and acceptor sequences was deleted and replaced by
 CC an engineered stabilising sequence. pF8C8SS is used in a method for
 CC prodn. of a desired heterologous protein in a eukaryotic host cell.
 CC The method comprises: transfecting a eukaryotic host cell with a vector
 CC encoding trans-activating protein; transfecting the host cell with
 CC an expression vector having a sequence of a ds DNA comprising, a
 CC stabilising sequence downstream of a promoter and upstream of a DNA
 CC -adenylation sequence upstream to a transcription termination site.
 CC The pref. host cells are human embryonic kidney cells (293) or JW2.
 CC The method provides useful quantities of a desired protein in a
 CC relatively short period of time without having to establish continuous
 CC prodn. The transient expression system optimises the interaction between
 CC specific vector components and certain trans-activating proteins.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 777 BP; 188 A; 205 C; 190 G; 194 T; 0 other;
 SQ Query Match 25.4%; Score 127; DB 10; Length 777;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATGCGG 76
    |||||
Db 343 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATGCGG 402
    |||||
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136
    |||||
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 462
    |||||
QY 137 CACCCCA 143
    |||||
Db 463 CACCCCA 469
    |||||

RESULT 15
AAD53783
ID AAD53783 standard; DNA; 884 BP.
XX
AC AAD53783;
XX
DT 28-MAY-2003 (first entry)
XX
XX CMV enhancer/Myeloproliferative sarcoma virus LTR promoter construct DNA.
DE
DE Transmembrane activator; calcium modulator; nephrotropic; antibacterial;
KW TAC1; tumour necrosis factor-like protein; ZTNF2; ZTNF4; immunoglobulin;
KW anaemia; gene therapy; cytostatic; antiinflammatory; immunosuppressive;
KW glomerulonephritis; asthma; bronchitis; graft rejection; septic shock;
KW dermatological; neuroprotective; cyclophilin ligand-interactor;
KW autoimmune disease; systemic lupus erythematosus; multiple sclerosis;
KW diabetes mellitus; rheumatoid arthritis; renal disease; inflammation;
KW Cytomegalovirus; CMV; chimeric; gene; ds.
XX
OS Chimeric - Cytomegalovirus.
OS Chimeric - Myeloproliferative sarcoma virus.
XX
FH Key Location/Qualifiers
FT misc_feature 1..407
FT FT /*tag= a
FT FT /note= "CMV enhancer"
FT FT misc_feature 408..884
FT FT /*tag= b
FT FT /note= "Myeloproliferative sarcoma virus LTR promoter"
XX
PN WO200294852-A2.
XX
XX 28-NOV-2002.
XX
XX 20-MAY-2002; 2002WO-US15910.
XX
XX 24-MAY-2001; 2001US-293343P.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Rixon MW, Gross JA;
XX
XX WPI; 2003-148455/14.
XX
XX Transmembrane activator and calcium modulator and cyclophilin
PT ligand-interactor (TAC1)-immunoglobulin fusion protein, for treating
PT cancer or diabetes, comprises a TAC1 receptor group and an
PT immunoglobulin group -
XX
XX Disclosure; Column 141; 71pp; English.
XX
XX The invention relates to fusion proteins comprising transmembrane
CC activator and calcium modulator and cyclophilin ligand-interactor (TAC1)
CC receptor group that binds tumour necrosis factor-like protein (ZTNF)2 or
CC ZTNF4, and an immunoglobulin group comprising a constant region of an
CC immunoglobulin. The invention is used to manufacture a medicament for
CC inhibiting the proliferation of tumour cells in a mammalian subject.
CC The composition comprising the fusion protein may also be used in
CC treating autoimmune diseases (e.g. systemic lupus erythematosus,
```

```
CC multiple sclerosis, diabetes mellitus, rheumatoid arthritis and asthma),
CC renal diseases (e.g. glomerulonephritis), bronchitis, inflammation,
CC graft rejection, anaemia and septic shock. The fusion proteins are
CC also used in gene therapy. The present sequence is Cytomegalovirus (CMV)
CC enhancer/Myeloproliferative sarcoma virus LTR promoter construct DNA
CC used in the invention.
```

XX Sequence 884 BP; 229 A; 236 C; 214 G; 205 T; 0 other;

Query Match 25.4%; Score 127; DB 25; Length 884;

Best Local Similarity 100.0%; Pred. No. 7,9e-59;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATGCGG 76

Db 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATGCGG 315

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136

Db 316 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 375

QY 137 CACCCCA 143

Db 376 CACCCCA 382

Search completed: February 15, 2004, 22:40:23

Job time : 165.871 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 10:22:55 ; Search time 36.7706 Seconds
(without alignments)
6001.851 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500
Sequence: 1 asgcttgacctatgcgact.....caoctgggtggatcgatc 500

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	25.4	753	US-08-459-493-1	Sequence 1, Appli
2	127	25.4	930	US-08-029-022-2	Sequence 2, Appli
3	127	25.4	930	US-08-029-022-4	Sequence 4, Appli
4	127	25.4	930	US-08-246-376-2	Sequence 2, Appli
5	127	25.4	930	US-08-246-376-4	Sequence 4, Appli
6	127	25.4	930	US-07-972-135-2	Sequence 2, Appli
7	127	25.4	930	US-07-972-135-4	Sequence 4, Appli
8	127	25.4	930	US-08-256-004-2	Sequence 2, Appli
9	127	25.4	930	US-09-006-841-2	Sequence 2, Appli
10	127	25.4	930	US-09-006-841-4	Sequence 4, Appli
11	127	25.4	930	PCT-US93-05366-2	Sequence 2, Appli
12	127	25.4	930	PCT-US93-05366-4	Sequence 4, Appli
13	127	25.4	1078	US-09-310-842-1	Sequence 1, Appli
14	127	25.4	1215	US-08-522-841-7	Sequence 7, Appli
15	127	25.4	1216	US-08-522-841-3	Sequence 3, Appli
16	127	25.4	1318	US-09-310-842-3	Sequence 3, Appli
17	127	25.4	1417	US-09-310-842-5	Sequence 5, Appli
18	127	25.4	1467	US-09-800-170-27	Sequence 27, Appli
19	127	25.4	1645	US-09-310-842-2	Sequence 2, Appli
20	127	25.4	1870	US-09-310-842-4	Sequence 4, Appli
21	127	25.4	2043	US-09-800-170-47	Sequence 47, Appli
22	127	25.4	3125	US-08-037-816A-13	Sequence 13, Appli
23	127	25.4	3125	US-08-530-146-13	Sequence 13, Appli
24	127	25.4	3547	US-09-340-798A-43	Sequence 43, Appli
25	127	25.4	3796	US-09-470-661A-32	Sequence 32, Appli
26	127	25.4	3853	US-08-801-092-5	Sequence 5, Appli
27	127	25.4	3853	US-09-315-113-5	Sequence 5, Appli

28	127	25.4	3925	3	US-09-011-745-9	Sequence 9, Appli
29	127	25.4	3944	1	US-07-678-408A-1	Sequence 1, Appli
30	127	25.4	3987	4	US-09-082-649B-83	Sequence 83, Appli
31	127	25.4	3987	4	US-09-082-649B-84	Sequence 84, Appli
32	127	25.4	4026	3	US-08-801-092-19	Sequence 19, Appli
33	127	25.4	4026	4	US-09-315-113-19	Sequence 19, Appli
34	127	25.4	4059	2	US-08-252-493C-4	Sequence 4, Appli
35	127	25.4	4059	3	US-09-276-197-4	Sequence 4, Appli
36	127	25.4	4059	4	US-08-487-283A-3	Sequence 3, Appli
37	127	25.4	4059	5	PCT-US95-0755A-4	Sequence 4, Appli
38	127	25.4	4059	5	PCT-US96-0561A-11	Sequence 11, Appli
39	127	25.4	4089	4	US-09-170-496D-287	Sequence 287, App
40	127	25.4	4089	4	US-09-170-496D-288	Sequence 288, App
41	127	25.4	4072	4	US-09-770-315-4	Sequence 4, Appli
42	127	25.4	4249	3	US-08-801-092-33	Sequence 33, Appli
43	127	25.4	4249	4	US-09-315-113-33	Sequence 33, Appli
44	127	25.4	4278	4	US-09-503-799-2	Sequence 2, Appli
45	127	25.4	4283	1	US-08-343-401A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-459-493-1
; Sequence 1, Application US/08459493
; Patent No. 6030638
; GENERAL INFORMATION:
; APPLICANT: Brigham, Kenneth
; APPLICANT: Conary, Jon T.
; APPLICANT: Canonico, Angelo
; APPLICANT: Meyrick, Barbara
; TITLE OF INVENTION: PLASMID FOR IN VIVO EXPRESSION
; TITLE OF INVENTION: PROTAGLANDIN SYNTHASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 S. Wacker Drive - Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/459,493
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/080,221
; FILING DATE: 21-JUNE-1993
; APPLICATION NUMBER: US 07/746,941
; FILING DATE: 19-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: VU9115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (313) 456-8000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: promoter

LOCATION: 1..763
US-08-459-493-1
Query Match 25.4%; Score 127; DB 3; Length 763;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATCGG 76
Db 492 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATCGG 551
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGGTCTCACGGGGATTTCACAGTCTC 136
Db 552 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGGTCTCACGGGGATTTCACAGTCTC 611
QY 137 CACCCCA 143
Db 612 CACCCCA 618
RESULT 2
US-08-029-022-2
; Sequence 2, Application US/08029022
; Patent No. 5641662
; GENERAL INFORMATION:
; APPLICANT: Debs, Robert J.
; APPLICANT: Zhu, Ning
; TITLE OF INVENTION: TRANSFECTION OF LUNG VIA AEROSOLIZED
; TITLE OF INVENTION: TRANSGENE DELIVERY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,022
; FILING DATE: 19930310
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 809,291
; FILING DATE: 17 DEC 91
; APPLICATION NUMBER: 972,135
; FILING DATE: 05 NOV 92
; APPLICATION NUMBER: PCT/US92/11008
; FILING DATE: 17 DEC 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter, Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/008001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-029-022-2
Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATCGG 76
Db 469 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGGTCTCACGGGGATTTCACAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGGTCTCACGGGGATTTCACAGTCTC 588
QY 137 CACCCCA 143
Db 589 CACCCCA 595
RESULT 3
US-08-029-022-4
; Sequence 4, Application US/08029022
; Patent No. 5641662
; GENERAL INFORMATION:
; APPLICANT: Debs, Robert J.
; APPLICANT: Zhu, Ning
; TITLE OF INVENTION: TRANSFECTION OF LUNG VIA AEROSOLIZED
; TITLE OF INVENTION: TRANSGENE DELIVERY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,022
; FILING DATE: 19930310
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 809,291
; FILING DATE: 17 DEC 91
; APPLICATION NUMBER: 972,135
; FILING DATE: 05 NOV 92
; APPLICATION NUMBER: PCT/US92/11008
; FILING DATE: 17 DEC 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter, Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/008001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-029-022-4
Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATCGG 76
Db 469 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGGTCTCACGGGGATTTCACAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGGTCTCACGGGGATTTCACAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 4

US-08-246-376-2
; Sequence 2, Application US/08246376
; Patent No. 5827703
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; TITLE OF INVENTION: GENE THERAPY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,376
; FILING DATE:

CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-246-376-2

Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACATTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACATTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTCCAAAGTCTC 588
QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 5

US-08-246-376-4
; Sequence 4, Application US/08246376
; Patent No. 5827703
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; TITLE OF INVENTION: GENE THERAPY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,376
; FILING DATE:

CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACATTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACATTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTCCAAAGTCTC 588
QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 6

US-07-972-135-2
; Sequence 2, Application US/07972135
; Patent No. 5858784
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: EXPRESSION OF CLONED
; TITLE OF INVENTION: GENES IN THE

;; TITLE OF INVENTION: LUNG BY AEROSOL - AND
;; TITLE OF INVENTION: LIPOSOME-BASED
;; TITLE OF INVENTION: DELIVERY
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson
;; STREET: 2200 Sand Hill Road,
;; CITY: Menlo Park
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 94025
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 50Z or
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/972,135
;; FILING DATE: December 17, 1991
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/809,291
;; FILING DATE: December 17, 1991
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barbara Rae-Venter
;; REGISTRATION NUMBER: 32,750
;; REFERENCE/DOCKET NUMBER: 05935/007US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 854-5277
;; TELEFAX: (415) 854-0875
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 930
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-07-972-135-4

Query Match 25.4%; Score 127; DB 2; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GAC TTCTCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GAC TTCTCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGG 528

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 588

Qy 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 8
US-08-256-004-2
; Sequence 2, Application US/08256004
; Patent No. 6001644
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; APPLICANT: Ning Zhu
; TITLE OF INVENTION: IN VIVO TRANSFECTION WITH A CFTR CODING
; TITLE OF INVENTION: SEQUENCE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: U.S.A.

;; TITLE OF INVENTION: LUNG BY AEROSOL - AND
;; TITLE OF INVENTION: LIPOSOME-BASED
;; TITLE OF INVENTION: DELIVERY
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson
;; STREET: 2200 Sand Hill Road,
;; CITY: Menlo Park
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 94025
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 50Z or
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/972,135
;; FILING DATE: December 17, 1991
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/809,291
;; FILING DATE: December 17, 1991
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barbara Rae-Venter
;; REGISTRATION NUMBER: 32,750
;; REFERENCE/DOCKET NUMBER: 05935/007US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 854-5277
;; TELEFAX: (415) 854-0875
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 930
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-07-972-135-2

Query Match 25.4%; Score 127; DB 2; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GAC TTCTCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GAC TTCTCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGG 528

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 588

Qy 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 7
US-07-972-135-4
; Sequence 4, Application US/07972135
; Patent No. 5858784
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: EXPRESSION OF CLONED
; TITLE OF INVENTION: GENES IN THE
; TITLE OF INVENTION: LUNG BY AEROSOL - AND
; TITLE OF INVENTION: LIPOSOME-BASED
; TITLE OF INVENTION: DELIVERY
; NUMBER OF SEQUENCES: 4

ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,004
FILING DATE: August 22, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/11004
FILING DATE: December 17, 1992
APPLICATION NUMBER: 07/972,135
FILING DATE: NO. 6001644ember 5, 1992
APPLICATION NUMBER: 07/927,200,
FILING DATE: August 6, 1992
APPLICATION NUMBER: 07/894,498
FILING DATE: June 4, 1992
APPLICATION NUMBER: 07/809,291
FILING DATE: December 17, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Neeley, Richard L.
REGISTRATION NUMBER: 30,092
REFERENCE/DOCKET NUMBER: UCSF-008/00US
TELEPHONE: (415) 843-5070
TELEFAX: (415) 857-0663
TELEX: 380816COOLEYPA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-004-2

Query Match 25.4%; Score 127; DB 3; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 76
Db 469 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 9
US-09-006-841-2
Sequence 2, Application US/09006841
Patent No. 6468798
GENERAL INFORMATION:
APPLICANT: Robert J. Debs
TITLE OF INVENTION: EXPRESSION OF CLONED
GENES IN THE
TITLE OF INVENTION: LUNG BY AEROSOL - AND
TITLE OF INVENTION: LIPOSOME-BASED
DELIVERY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road,
SUITE: Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/972,135
FILING DATE: No. 6468798ember 5, 1992
APPLICATION NUMBER: 07/809,291
FILING DATE: December 17, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Barbara Rae-Venter
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: 05935/007US1
TELEPHONE: (415) 854-5277
TELEFAX: (415) 854-0875
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-006-841-2

Query Match 25.4%; Score 127; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 76
Db 469 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 10
US-09-006-841-4
Sequence 4, Application US/09006841
Patent No. 6468798
GENERAL INFORMATION:
APPLICANT: Robert J. Debs
TITLE OF INVENTION: EXPRESSION OF CLONED
GENES IN THE
TITLE OF INVENTION: LUNG BY AEROSOL - AND
TITLE OF INVENTION: LIPOSOME-BASED
DELIVERY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road,
SUITE: Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006.841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/972.135
; FILING DATE: No. 6468798ember 5, 1992
; APPLICATION NUMBER: 07/809.291
; FILING DATE: December 17, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/007US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-006-841-4

Query Match 25.4%; Score 127; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACGGGGATTTCCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACGGGGATTTCCAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595
```

```

RESULT 11
PCT-US93-05366-2
; Sequence 2, Application PC/TUS9305366
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05366
; FILING DATE: 04-JUN-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
```

```

; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
PCT-US93-05366-2

Query Match 25.4%; Score 127; DB 5; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACGGGGATTTCCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACGGGGATTTCCAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 12
PCT-US93-05366-4
; Sequence 4, Application PC/TUS9305366
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05366
; FILING DATE: 04-JUN-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
```

TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US93-05386-4

Query Match 25.4%; Score 127; DB 5; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 76
DB 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
DB 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 588

QY 137 CACCCCA 143
DB 589 CACCCCA 595

RESULT 13
US-09-310-842-1
Sequence 1, Application US/09310842A
Patent No. 6451593
GENERAL INFORMATION:
APPLICANT: Wittig, Prof. Burghardt
APPLICANT: Jungmans, Claas
TITLE OF INVENTION: Design Principle for Constructing Expression Constructs for Gene
TITLE OF INVENTION: Therapy
FILE REFERENCE: XI 597/99
CURRENT APPLICATION NUMBER: US/09/310,842A
CURRENT FILING DATE: 1999-05-12
EARLIER APPLICATION NUMBER: DE 196 48 625.4
EARLIER FILING DATE: 1996-11-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 1078
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: gene (1078)
LOCATION: (1)..(1078)
OTHER INFORMATION: bcr3=abl2; Oligo DNA Dumbbell
FEATURE:
NAME/KEY: misc binding
LOCATION: (1)..(2)
OTHER INFORMATION: intramolecular binding site; the T-nucleotides at position 1 to 1077 can be modified with amino or caroxy features
FEATURE:
NAME/KEY: misc binding
LOCATION: (1077)..(1078)
OTHER INFORMATION: intramolecular binding site; the T-nucleotides at position 1077 to 1078 can be modified with amino or caroxy features
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Strandedness: both; nucleic acid (linear), hypothetical: No. 6451593 anti-sense: No
US-09-310-842-1

Query Match 25.4%; Score 127; DB 4; Length 1078;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 76
DB 352 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 411

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
DB 412 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 471

QY 137 CACCCCA 143
DB 472 CACCCCA 478

RESULT 14
US-08-522-841-7
Sequence 7, Application US/08522841
Patent No. 5919652
GENERAL INFORMATION:
APPLICANT: Pang, Shen
APPLICANT: Belledgrun, Arie S.
TITLE OF INVENTION: No. 5919652el Nucleic Acid Molecules Comprising
TITLE OF INVENTION: the Prostate Specific Antigen (PSA) Promoter and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant and Gould
STREET: 11150 Santa Monica Blvd., Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/522,841
FILING DATE: 01-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30435.26US11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1215 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-522-841-7

Query Match 25.4%; Score 127; DB 2; Length 1215;
Best Local Similarity 100.0%; Pred. No. 2.4e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 76
DB 333 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 392

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
DB 393 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 452

QY 137 CACCCCA 143
DB 453 CACCCCA 459

RESULT 15
US-08-522-841-3
Sequence 3, Application US/08522841.
Patent No. 5919652

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 16:21:35 ; Search time 210.032 Seconds
(without alignments)
8769.219 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500

Sequence: 1 aagcttgacattatgcagct.....cacctgggtggatgc 500

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/PCTUS_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	500	100.0	5283	13	US-09-921-143-36
2	127	25.4	351	13	US-10-268-927-35
3	127	25.4	412	11	US-09-911-020A-1
4	127	25.4	544	13	US-10-268-927-33
5	127	25.4	600	9	US-09-895-828-425
6	127	25.4	600	15	US-10-114-666-425
7	127	25.4	619	10	US-09-965-703-68
8	127	25.4	645	13	US-10-437-830-3
9	127	25.4	851	12	US-10-465-156-1
10	127	25.4	884	15	US-10-152-363A-69
11	127	25.4	930	15	US-10-206-747-2
12	127	25.4	930	15	US-10-206-747-4
13	127	25.4	987	13	US-09-845-416-33
14	127	25.4	996	8	US-08-900-220C-25
15	127	25.4	996	10	US-09-151-999-25

16	127	25.4	996	11	US-09-187-387-25	Sequence 25, Appl
17	127	25.4	996	11	US-09-827-110-25	Sequence 25, Appl
18	127	25.4	996	13	US-09-238-243-25	Sequence 25, Appl
19	127	25.4	1070	15	US-10-205-534-6	Sequence 6, Appl
20	127	25.4	1078	15	US-10-228-811-1	Sequence 1, Appl
21	127	25.4	1104	10	US-09-845-899A-6	Sequence 6, Appl
22	127	25.4	1318	15	US-10-228-811-3	Sequence 3, Appl
23	127	25.4	1417	15	US-10-228-811-5	Sequence 5, Appl
24	127	25.4	1567	12	US-10-431-121A-3	Sequence 3, Appl
25	127	25.4	1645	15	US-10-228-811-2	Sequence 2, Appl
26	127	25.4	1715	9	US-09-886-942-10	Sequence 10, Appl
27	127	25.4	1767	9	US-09-886-942-8	Sequence 8, Appl
28	127	25.4	1767	9	US-09-886-942-19	Sequence 19, Appl
29	127	25.4	1787	12	US-10-364-049-900	Sequence 900, App
30	127	25.4	1813	9	US-03-794-975-8	Sequence 8, Appl
31	127	25.4	1870	15	US-10-228-811-4	Sequence 4, Appl
32	127	25.4	2335	15	US-10-078-167-2	Sequence 2, Appl
33	127	25.4	3360	15	US-10-023-969-4	Sequence 4, Appl
34	127	25.4	3426	13	US-10-136-837-2	Sequence 2, Appl
35	127	25.4	3487	15	US-10-303-683-16	Sequence 16, Appl
36	127	25.4	3547	12	US-10-369-121-43	Sequence 43, Appl
37	127	25.4	3589	13	US-10-136-837-4	Sequence 4, Appl
38	127	25.4	3589	15	US-10-234-406-1	Sequence 1, Appl
39	127	25.4	3609	15	US-10-234-406-3	Sequence 3, Appl
40	127	25.4	3610	11	US-09-194-949-1	Sequence 1, Appl
41	127	25.4	3788	9	US-09-392-462-1	Sequence 1, Appl
42	127	25.4	3788	9	US-09-540-991-1	Sequence 1, Appl
43	127	25.4	3788	13	US-10-291-041-1	Sequence 1, Appl
44	127	25.4	3796	13	US-10-241-332-32	Sequence 32, Appl
45	127	25.4	3796	14	US-10-127-391-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-921-143-36
; Sequence 36, Application US/09921143
; Publication No. US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P6
; CURRENT APPLICATION NUMBER: US/09/921,143
; CURRENT FILING DATE: 2001-08-03
; PRIOR FILING DATE: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent version 3.0
; SEQ ID NO 36
; LENGTH: 5283
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-36

Query Match	100.0%	Score	500	DB	13	Length	5283
Best Local Similarity	100.0%	Pred. No.	8.8e-272				
Matches	500	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	AAGCTTGACCTTATGGGACTTTCTCTACTTTGGCAGTACATCTACGTATTAGTCATCGCTAT	60				
Db	1	AAGCTTGACCTTATGGGACTTTCTCTACTTTGGCAGTACATCTACGTATTAGTCATCGCTAT	60				
Qy	61	TACATGGTGATCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACATCACG	120				
Db	61	TACATGGTGATCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACATCACG	120				
Qy	121	GGGATTTCCAAAGTCTCCACCCCACTGACGCTCAATGGGAGTTGTTTGGACCAAAATCA	180				
Db	121	GGGATTTCCAAAGTCTCCACCCCACTGACGCTCAATGGGAGTTGTTTGGACCAAAATCA	180				
Qy	181	ACGAGATTTCCAAATGTCGTAACACTCCGCCCATTCACGCAATGGCGGTAGGCG	240				
Db	181	ACGAGATTTCCAAATGTCGTAACACTCCGCCCATTCACGCAATGGCGGTAGGCG	240				

Db 181 ACAGACTTCCAAAATGCTGTAACAACACTCGCCGCCCAATTGACGCAAAATGGCGGTAGGCG 240
Qy 241 RACATGCTTATGTAACGGTGTAGTACCAACATGCTTATAAGGAGAGAAAAGCACCGTG 300
Db 241 RACATGCTTATGTAACGGTGTAGTACCAACATGCTTATAAGGAGAGAAAAGCACCGTG 300
Qy 301 CATGCCGATTGGTGGAGTAAGGTGCTGATGATCGTGGTATGATCGTGGCTTTGTAGGAAG 360
Db 301 CATGCCGATTGGTGGAGTAAGGTGCTGATGATCGTGGTATGATCGTGGCTTTGTAGGAAG 360
Qy 361 GCAACAGACGGGTCTAAACACGATTTGAGCAGACCACTCAATTCGCGCATTCAGAGATATT 420
Db 361 GCAACAGACGGGTCTAAACACGATTTGAGCAGACCACTCAATTCGCGCATTCAGAGATATT 420
Qy 421 GTATTTAAGTGGCCCGCTGATACAAATAAAGCCCAATTTGACCAATTCACCAATTTGGTGTG 480
Db 421 GTATTTAAGTGGCCCGCTGATACAAATAAAGCCCAATTTGACCAATTCACCAATTTGGTGTG 480
Qy 481 CACCTGGGTGGATCGATC 500
Db 481 CACCTGGGTGGATCGATC 500

RESULT 2

US-10-268-927-35

; Sequence 35, Application US/10268927

; Publication No. US20030175693A1

; GENERAL INFORMATION:

; APPLICANT: Main-Hobson, Simon

; APPLICANT: Blancou, Philippe

; APPLICANT: Chenciner, Nicole

; TITLE OF INVENTION: HIV RECOMBINANT VACCINE

; FILE REFERENCE: 03495.0239-00000

; CURRENT APPLICATION NUMBER: US/10/268,927

; CURRENT FILING DATE: 2002-10-11

; PRIOR APPLICATION NUMBER: 60/328,449

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 35

; LENGTH: 351

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: recombinant promoter

; US-10-268-927-35

Query Match 25.4%; Score 127; DB 13; Length 351;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 76
Db 74 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 133
Qy 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 134 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 193
Qy 137 CACCCCA 143
Db 194 CACCCCA 200

RESULT 3

US-09-911-020A-1

; Sequence 1, Application US/09911020A

; Publication No. US20030040100A1

; GENERAL INFORMATION:

; APPLICANT: BROUGH, DOUGLAS E

; APPLICANT: KOVESDI, IMRE

; TITLE OF INVENTION: CELL FOR THE PROPAGATION OF ADENOVIRAL VECTORS

; FILE REFERENCE: 209459

; CURRENT APPLICATION NUMBER: US/09/911,020A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Cytomegalovirus
; US-09-911-020A-1

Query Match 25.4%; Score 127; DB 11; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 76
Db 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 315
Qy 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 316 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 375
Qy 137 CACCCCA 143
Db 376 CACCCCA 382

RESULT 4

US-10-268-927-33

; Sequence 33, Application US/10268927

; Publication No. US20030175693A1

; GENERAL INFORMATION:

; APPLICANT: Main-Hobson, Simon

; APPLICANT: Blancou, Philippe

; APPLICANT: Chenciner, Nicole

; TITLE OF INVENTION: HIV RECOMBINANT VACCINE

; FILE REFERENCE: 03495.0239-00000

; CURRENT APPLICATION NUMBER: US/10/268,927

; CURRENT FILING DATE: 2002-10-11

; PRIOR APPLICATION NUMBER: 60/328,449

; PRIOR FILING DATE: 2001-10-12

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 33

; LENGTH: 544

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: recombinant promoter

; US-10-268-927-33

Query Match 25.4%; Score 127; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 76
Db 267 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 326
Qy 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 327 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 386
Qy 137 CACCCCA 143
Db 387 CACCCCA 393

RESULT 5

US-09-895-828-425

; Sequence 425, Application US/09895828

; Patent No. US2002009012A1

; GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: McNeill, Patricia D.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Carter, Darrick
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.539
CURRENT APPLICATION NUMBER: US/09/895,828
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 473
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 425
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-828-425

Query Match 25.4%; Score 127; DB 9; Length 600;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 278 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATCGG 337

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 338 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 397

QY 137 CACCCCA 143
Db 398 CACCCCA 404

RESULT 6
US-10-114-666-425
Sequence 425, Application US/10114666
Publication No. US20030103994A1
GENERAL INFORMATION:
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.539C1
CURRENT APPLICATION NUMBER: US/10/114,666
CURRENT FILING DATE: 2002-04-01
NUMBER OF SEQ ID NOS: 479
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 425
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-666-425

Query Match 25.4%; Score 127; DB 15; Length 600;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 278 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATCGG 337

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 338 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 397

QY 137 CACCCCA 143
Db 398 CACCCCA 404

RESULT 7
US-09-965-703-68
Sequence 68, Application US/09965703
Patent No. US20020119521A1
GENERAL INFORMATION:
APPLICANT: Rohm and Haas Company
APPLICANT: Palli, Subba Reddy
APPLICANT: Kapitskaya, Marianna zinovjevna
APPLICANT: Cress, Dean Ervin
TITLE OF INVENTION: No. US20020119521A1el Bodysone Receptor-Based Inducible Gene E
FILE REFERENCE: A01020B
CURRENT APPLICATION NUMBER: US/09/965,703
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/191,355
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/269,799
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: PCT/US01/09050
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 68
LENGTH: 619
TYPE: DNA
ORGANISM: Cytomegalovirus
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: No. US20020119521A1el Sequence
US-09-965-703-68

Query Match 25.4%; Score 127; DB 10; Length 619;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 254 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATCGG 313

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 314 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 373

QY 137 CACCCCA 143
Db 374 CACCCCA 380

RESULT 8
US-10-437-830-3
Sequence 3, Application US/10437830
Publication No. US20030180949A1
GENERAL INFORMATION:
APPLICANT: JOHN LEVY
TITLE OF INVENTION: AVOIDANCE OF UNDESIRABLE REPLICATION
TITLE OF INVENTION: INTERMEDIATES IN PLASMID PROPAGATION
FILE REFERENCE: CT1MM.014DV1
CURRENT APPLICATION NUMBER: US/10/437,830
CURRENT FILING DATE: 2003-05-13
PRIOR APPLICATION NUMBER: 09/715,835
PRIOR FILING DATE: 2000-11-16
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 645
TYPE: DNA
ORGANISM: CYTOMEGALOVIRUS
US-10-437-830-3

Query Match 25.4%; Score 127; DB 13; Length 645;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTACCATGCTGTATCGCGG 76
DB 431 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTACCATGCTGTATCGCGG 490
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
DB 491 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 550
QY 137 CACCCCA 143
DB 551 CACCCCA 557

RESULT 9
US-10-465-156-1
; Sequence 1, Application US/10465156
; Publication No. US20030232414A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Margaret Dow
; TITLE OF INVENTION: HYBRID VECTOR HAVING A CYTOMEGALOVIRUS
; TITLE OF INVENTION: ENHANCER AND MYELOPROLIFERATIVE SARCOMA VIRUS PROMOTER
; FILE REFERENCE: 02-15
; CURRENT APPLICATION NUMBER: US/10/465,156
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/389,612
; PRIOR FILING DATE: 2002-06-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hybrid cytomegalovirus and myeloproliferative
; OTHER INFORMATION: sarcoma virus regulatory sequence
US-10-465-156-1

Query Match 25.4%; Score 127; DB 12; Length 851;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTACCATGCTGTATCGCGG 76
DB 223 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTACCATGCTGTATCGCGG 282
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
DB 283 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 342
QY 137 CACCCCA 143
DB 343 CACCCCA 349

RESULT 10
US-10-152-363A-69
; Sequence 69, Application US/10152363A
; Publication No. US20030103986A1
; GENERAL INFORMATION:
; APPLICANT: Rixon, Mark W.
; APPLICANT: Gross, Jane A.
; TITLE OF INVENTION: TACT-Immunoglobulin Fusion Proteins
; FILE REFERENCE: 01-20
; CURRENT APPLICATION NUMBER: US/10/152,363A
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/293,343
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: CMV enhancer/MPSV LTR promoter construct.
US-10-152-363A-69

Query Match 25.4%; Score 127; DB 15; Length 884;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTACCATGCTGTATCGCGG 76
DB 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTACCATGCTGTATCGCGG 315
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136
DB 316 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 375
QY 137 CACCCCA 143
DB 376 CACCCCA 382

RESULT 11
US-10-206-747-2
; Sequence 2, Application US/10206747
; Publication No. US20030109475A1
; GENERAL INFORMATION:
; APPLICANT: (i) APPLICANT: Zhu, Ning
; (i) APPLICANT: Zhu, Ning
; Zhu, Ning
; TITLE OF INVENTION: Methods and Compositions for In Vivo
; Gene Therapy
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,747
; FILING DATE: 11-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/809,292
; FILING DATE: 17-DEC-1991
; APPLICATION NUMBER: US 07/894,498
; FILING DATE: 04-JUN-1992
; APPLICATION NUMBER: US 07/927,200
; FILING DATE: 06-AUG-1992
; APPLICATION NUMBER: US 07/972,135
; FILING DATE: 05-NOV-1992
; APPLICATION NUMBER: US 07/992,687
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US92/11004
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US93/05366
; FILING DATE: 04-JUN-1993
; APPLICATION NUMBER: US 08/464,899
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: US 09/132,391
; FILING DATE: 08-NOV-2003
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J.
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 023070-064833US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
```



```
;
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 930 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-206-747-2

Query Match      25.4%; Score 127; DB 15; Length 930;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 12
US-10-206-747-4
; Sequence 4, Application US/10206747
; Publication No. US20030109475A1
; GENERAL INFORMATION:
; APPLICANT: (i) APPLICANT: Zhu, Ning
;              (ii) APPLICANT: Zhu, Ning
;              Zhu, Ning
; TITLE OF INVENTION: Methods and Compositions for In Vivo
;                 Gene Therapy
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,747
; FILING DATE: 11-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/809,292
; FILING DATE: 17-DEC-1991
; APPLICATION NUMBER: US 07/894,498
; FILING DATE: 04-JUN-1992
; APPLICATION NUMBER: US 07/927,200
; FILING DATE: 06-AUG-1992
; APPLICATION NUMBER: US 07/972,135
; FILING DATE: 05-NOV-1992
; APPLICATION NUMBER: US 07/992,687
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US92/11004
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US93/05366
; FILING DATE: 04-JUN-1993
; APPLICATION NUMBER: US 08/464,899
; FILING DATE: 22-JAN-1996

;
; APPLICATION NUMBER: US 09/132,391
; FILING DATE: 08-No. US20030109475A1-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J.
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 023070-064833US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 930 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-206-747-4

Query Match      25.4%; Score 127; DB 15; Length 930;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 13
US-09-845-416-33
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-33

Query Match      25.4%; Score 127; DB 13; Length 987;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 417 GACTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTATTACCATGGTGATCGG 476

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 477 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 536

QY 137 CACCCCA 143
Db 537 CACCCCA 543
```

RESULT 14

US-08-900-220C-25
; Sequence 25, Application US/08900220C
; Publication No. US20020045206A1

GENERAL INFORMATION:
; APPLICANT: Miao, Ningning

Wang, Monica
Mahanthappa, Nagesh K.
Pang, Kevin
Jin, Ping

TITLE OF INVENTION: Method of Treating Dopaminergic and

GABA-nergic Disorders

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: POLEY, HOAG & ELIOT LLP

STREET: ONE POST OFFICE SQUARE

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/900,220C

FILING DATE: 24-Jul-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: ONV-044.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 996 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: nucleic acid

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-08-900-220C-25

Query Match 25.4%; Score 127; DB 8; Length 996;

Best Local Similarity 100.0%; Pred. No. 5.2e-61;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGCATCGG 76

Db 593 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGCATCGG 652

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCAGCGGGATTTCCAAAGTCTC 136

Db 653 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCAGCGGGATTTCCAAAGTCTC 712

Qy 137 CACCCCA 143

Db 713 CACCCCA 719

RESULT 15

US-09-151-999-25

; Sequence 25, Application US/09151999

; Patent No. US20020151460A1

GENERAL INFORMATION:

; APPLICANT: Wang, Elizabeth

; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE

; TITLE OF INVENTION: POLYPEPTIDES, AND FORMULATIONS AND USES RELATED THERETO

; FILE REFERENCE: ONV-031.02

; CURRENT APPLICATION NUMBER: US/09/151,999

; CURRENT FILING DATE: 1998-08-11

; EARLIER APPLICATION NUMBER: 08/955,552

; EARLIER FILING DATE: 1997-10-20

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 25

; LENGTH: 996

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: gene

; OTHER INFORMATION: activation construct

US-09-151-999-25

Query Match 25.4%; Score 127; DB 10; Length 996;

Best Local Similarity 100.0%; Pred. No. 5.2e-61;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGCATCGG 76

Db 593 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGCATCGG 652

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCAGCGGGATTTCCAAAGTCTC 136

Db 653 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCAGCGGGATTTCCAAAGTCTC 712

Qy 137 CACCCCA 143

Db 713 CACCCCA 719

Search completed: February 15, 2004, 22:26:45

Job time : 213.032 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 21:50:21 ; Search time 1316.95 Seconds
(without alignments)
9227.587 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500

Sequence: 1 aagcttgacattatgcgact.....cacctgggtggatcgatc 500

Scoring table: OLIGO NUC

Gapop_60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_rtr:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	25.4	187	12	BI489763
2	127	25.4	191	12	BI489620
3	127	25.4	193	12	BI490268
4	127	25.4	193	12	BI490463

c	5	127	25.4	480	29	B2996651
c	6	127	25.4	500	29	B2680401
c	7	127	25.4	610	28	B68191
c	8	127	25.4	619	9	AV645385
c	9	127	25.4	634	9	AV645390
c	10	127	25.4	637	9	AV645392
c	11	127	25.4	640	9	AV645393
c	12	127	25.4	648	9	AV681492
c	13	127	25.4	655	9	AV681500
c	14	127	25.4	657	9	AV681507
c	15	127	25.4	661	9	AV681509
c	16	127	25.4	663	9	AV681529
c	17	127	25.4	681	9	AV681512
c	18	127	25.4	681	9	AV681513
c	19	127	25.4	684	9	AV681526
c	20	127	25.4	694	9	AV645336
c	21	127	25.4	733	9	AV681507
c	22	127	25.4	744	9	AV681465
c	23	127	25.4	764	29	B2817918
c	24	127	25.4	768	9	AV681525
c	25	127	25.4	772	9	AV681495
c	26	127	25.4	776	9	AV681505
c	27	127	25.4	783	9	AV681502
c	28	127	25.4	786	9	AV681504
c	29	127	25.4	802	9	AV681486
c	30	127	25.4	803	9	AV681483
c	31	127	25.4	806	9	AV681464
c	32	127	25.4	807	9	AV681462
c	33	127	25.4	810	9	AV681487
c	34	127	25.4	811	9	AV681506
c	35	127	25.4	816	9	AV681523
c	36	127	25.4	821	9	AV681519
c	37	127	25.4	834	29	B2667422
c	38	127	25.4	888	13	BX330112
c	39	127	25.4	905	13	BX459694
c	40	127	25.4	919	13	BX372036
c	41	127	25.4	920	13	BX393430
c	42	127	25.4	932	13	BX393879
c	43	127	25.4	948	13	BX364612
c	44	127	25.4	956	13	BX359183
c	45	127	25.4	963	13	BX342541

ALIGNMENTS

RESULT 1
BI489763
LOCUS 187 bp mRNA linear EST 28-AUG-2001
DEFINITION 603032280F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173506 5', mRNA sequence.
ACCESSION BI489763
VERSION BI489763.1 GI:15328991
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 187)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing By: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLNL1431 row: p column: 19
High quality sequence stop: 187.

BI489763 187 bp mRNA linear EST 28-AUG-2001
603032280F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173506 5', mRNA sequence.
ACCESSION BI489763
VERSION BI489763.1 GI:15328991
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 187)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing By: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLNL1431 row: p column: 19
High quality sequence stop: 187.

```

FEATURES
  source
    Location/Qualifiers
      1..187
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5173506"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_115"
        /note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPOR6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
      38 a 44 c 48 g 57 t
BASE COUNT
ORIGIN
  Query Match 25.4%; Score 127; DB 12; Length 187;
  Best Local Similarity 100.0%; Pred. No. 5.5e-58;
  Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 17 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGGCGG 76
  Db 39 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGGCGG 98
  QY 77 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
  Db 99 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 158
  QY 137 CACCCCA 143
  Db 159 CACCCCA 165
RESULT 2
BI489620
LOCUS 603032080F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173483 5',
DEFINITION mRNA sequence.
ACCESSION BI489620
VERSION BI489620.1 GI:15328848
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  1 (bases 1 to 191)
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  NIH-MGC http://mgc.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM11431 row: 0 column: 20
  High quality sequence stop: 191.
FEATURES
  source
    Location/Qualifiers
      1..191
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5173483"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_115"
        /note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPOR6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
      38 a 44 c 48 g 57 t
BASE COUNT
ORIGIN
  Query Match 25.4%; Score 127; DB 12; Length 187;
  Best Local Similarity 100.0%; Pred. No. 5.5e-58;
  Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 17 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGGCGG 76
  Db 39 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGGCGG 98
  QY 77 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
  Db 99 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 158
  QY 137 CACCCCA 143
  Db 159 CACCCCA 165
RESULT 2
BI489620
LOCUS 603032080F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173483 3',
DEFINITION mRNA sequence.
ACCESSION BI490268
VERSION BI490268.1 GI:15329483
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  1 (bases 1 to 193)
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  NIH-MGC http://mgc.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM11431 row: 0 column: 20
  High quality sequence stop: 193.
FEATURES
  source
    Location/Qualifiers
      1..193
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5173483"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_115"
        /note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPOR6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.

```

Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

```

BASE COUNT      59 a      50 c      46 g      38 t
ORIGIN

Query Match      25.4%; Score 127; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.6e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTGTATGGCG 76
Db 154 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTGTATGGCG 95

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 94 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136

QY 137 CACCCCA 143
Db 34 CACCCCA 28

RESULT 4
BI490463/c
LOCUS      BI490463
DEFINITION      193 bp mRNA linear EST 28-AUG-2001
ACCESSION      BI490463
VERSION      BI490463.1 GI:15329691
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 193)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11431 row: p column: 19
High quality sequence stop: 193.
FEATURES
source      1..193
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173506"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
PCMV-SF0R16; Site_1: Not; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      59 a      50 c      46 g      38 t
ORIGIN

Query Match      25.4%; Score 127; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.6e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTGTATGGCG 76
Db 154 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTGTATGGCG 95

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 94 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136

QY 137 CACCCCA 143
Db 34 CACCCCA 28

RESULT 5
BZ996651/c
LOCUS      BZ996651
DEFINITION      480 bp DNA linear GSS 25-MAR-2003
ACCESSION      BZ996651
VERSION      BZ996651.1 GI:29240068
KEYWORDS      GSS.
SOURCE      Zea mays
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 480)
AUTHORS      Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
TITLE      Maize Genomics Consortium
JOURNAL      Unpublished
COMMENT      Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source      1..480
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMWETA074N10"
/clone_lib="ZM 0.6-1.0 KB"
/notes="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

BASE COUNT      133 a      108 c      118 g      121 t
ORIGIN

Query Match      25.4%; Score 127; DB 29; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.5e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTGTATGGCG 76
Db 164 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTGTATGGCG 105

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 104 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 45

QY 137 CACCCCA 143
Db 44 CACCCCA 38

RESULT 6
BZ680401
LOCUS      BZ680401
DEFINITION      500 bp DNA linear GSS 05-FEB-2003
ACCESSION      BZ680401
VERSION      BZ680401.1 GI:15329691
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 193)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11431 row: p column: 19
High quality sequence stop: 193.
FEATURES
source      1..193
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173506"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
PCMV-SF0R16; Site_1: Not; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      59 a      50 c      46 g      38 t
ORIGIN

Query Match      25.4%; Score 127; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.6e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DEFINITION PUBIS38TD ZM.0.5.1.0.KB Zea mays genomic clone ZMWBTa064H03,
Genomic survey sequence.
ACCESSION BZ680401
VERSION BZ680401.1 GI:28235262
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1. (bases 1 to 500)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
Maize Genomics Consortium
Unpublished
Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
    source
    1..500
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMWBTa064H03"
    /clone_lib="ZM.0.5.1.0.KB"
    /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
    Cot selected genomic DNA library"
BASE COUNT 123 a 129 c 120 g 128 t
ORIGIN
Query Match 25.4%; Score 127; DB 29; Length 500;
Best Local Similarity 100.0%; Pred. No. 6.5e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCCTACTGTCAGTACATCTACGTTAGTCGCTATTACCATGTCGTCGG 76
Db 256 GACTTTCCTACTGTCAGTACATCTACGTTAGTCGCTATTACCATGTCGTCGG 315
QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTGACTCAGCGGGATTCCAGTCTC 136
Db 316 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTGACTCAGCGGGATTCCAGTCTC 375
QY 137 CACCCCA 143
Db 376 CACCCCA 382
RESULT 7
B68191/c
LOCUS B68191
DEFINITION CIT978SK-A-492C12.TV CIT978SK Homo sapiens genomic clone A-492C12,
genomic survey sequence.
ACCESSION B68191
VERSION B68191.1 GI:2666901
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 610)
Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden
K., Berry, K., Granger, D., Suh, E., Wible, C., Kim, U.-J., Shizuya, H.,
Simon, M. and Venter, J.C.
Use of a human BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished
Other_GSSs: CIT978SK-A-492C12.TP

```

```

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: 27
Class: BAC ends.
FEATURES
    source
    1..610
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="GDB:726476"
    /db_xref="taxon:9606"
    /clone="A-492C12"
    /sex="Female"
    /cell_type="Fibroblast"
    /clone_lib="CIT978SK"
    /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
    Caltech Human BAC Library A"
BASE COUNT 146 a 159 c 165 g 140 t
ORIGIN
Query Match 25.4%; Score 127; DB 28; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCCTACTGTCAGTACATCTACGTTAGTCGCTATTACCATGTCGTCGG 76
Db 508 GACTTTCCTACTGTCAGTACATCTACGTTAGTCGCTATTACCATGTCGTCGG 449
QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTGACTCAGCGGGATTCCAGTCTC 136
Db 448 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTGACTCAGCGGGATTCCAGTCTC 389
QY 137 CACCCCA 143
Db 388 CACCCCA 382
RESULT 8
AV645385/c
LOCUS AV645385
DEFINITION AV645385 GLA Homo sapiens cDNA clone GLAAG02 3', mRNA sequence.
ACCESSION AV645385
VERSION AV645385.1 GI:9866399
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 619)
Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z. and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
MEDLINE
PUBMED 11752456
COMMENT
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

```

This clone is available at CHGC in Shanghai.

```

FEATURES
  source
    1..619
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="GLAAG02"
      /tissue_type="corresponding non cancerous liver tissue"
      /dev_stage="Adult"
      /lab_host="SOLR"
      /clone_lib="GLA"
      /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
      XhoI"

BASE COUNT  147 a  164 c  159 g  145 t      4 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 619;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
    |||||
Db   385 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 326

QY  77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 136
    |||||
Db   325 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 266

QY  137 CACCCCA 143
    |||||
Db   265 CACCCCA 259

RESULT 9
AV645390/c
LOCUS      AV645390 634 bp mRNA linear EST 15-JAN-2002
DEFINITION AV645390 GLA Homo sapiens cDNA clone GLAAG07 3', mRNA sequence.
ACCESSION  AV645390
VERSION     AV645390.1 GI:9866404
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 634)
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
PUBMED     11752456
COMMENT    Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
  source
    1..634
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="GLAAG07"
      /tissue_type="corresponding non cancerous liver tissue"
      /dev_stage="Adult"
      /lab_host="SOLR"
      /clone_lib="GLA"

BASE COUNT  157 a  157 c  154 g  157 t  12 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 637;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
    |||||

```



```

Db 385 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATCGG 326
Qy 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGTTTGACTCACGGGGATTTTCCAAAGTCTC 136
Db 325 TTTTGGCAGTACATCAATGGGCGTGATAGCGTTTGACTCACGGGGATTTTCCAAAGTCTC 265
Qy 137 CACCCCA 143
Db 265 CACCCCA 259

```

```

RESULT 11
AV645393/c
LOCUS AV645393 640 bp mRNA linear EST 15-JAN-2002
DEFINITION AV645393 GLA Homo sapiens cDNA clone GLAAG10 3', mRNA sequence.
ACCESSION AV645393
VERSION AV645393.1 GI:9866407
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 640)
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z., and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

```

```

FEATURES
source
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLAAG10"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLA"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 152 a 162 c 164 g 153 t 9 others
ORIGIN
Query Match 25.4%; Score 127; DB 9; Length 640;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATCGG 76
Db 407 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATCGG 348
Qy 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGTTTGACTCACGGGGATTTTCCAAAGTCTC 136
Db 347 TTTTGGCAGTACATCAATGGGCGTGATAGCGTTTGACTCACGGGGATTTTCCAAAGTCTC 288
Qy 137 CACCCCA 143
Db 287 CACCCCA 281

```

```

RESULT 12
AV681492/c
LOCUS AV681492 648 bp mRNA linear EST 16-JAN-2002
DEFINITION AV681492 GLA Homo sapiens cDNA clone GKAAAD01 5', mRNA sequence.
ACCESSION AV681492
VERSION AV681492.1 GI:10283355
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 648)
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z., and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

```

```

FEATURES
source
1..648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKAAAD01"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKA"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 152 a 166 c 166 g 158 t 6 others
ORIGIN
Query Match 25.4%; Score 127; DB 9; Length 648;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATCGG 76
Db 387 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATCGG 328
Qy 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGTTTGACTCACGGGGATTTTCCAAAGTCTC 136
Db 327 TTTTGGCAGTACATCAATGGGCGTGATAGCGTTTGACTCACGGGGATTTTCCAAAGTCTC 268
Qy 137 CACCCCA 143
Db 267 CACCCCA 261

```

```

RESULT 13
AV681500
LOCUS AV681500 655 bp mRNA linear EST 16-JAN-2002
DEFINITION AV681500 GLA Homo sapiens cDNA clone GKAAAD08 5', mRNA sequence.
ACCESSION AV681500
VERSION AV681500.1 GI:10283363
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE
21625106
PUBMED
11752456
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
source
1..655
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="GKAPD08"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKA"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT
169 a 174 c 163 g 148 t 1 others
ORIGIN
Query Match 25.4%; Score 127; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 315
Qy 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 316 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 375
Qy 137 CACCCCA 143
Db 376 CACCCCA 382
RESULT 14
A1815377/c
LOCUS
DEFINITION
657 bp mRNA linear EST 09-JUL-1999
Danio rerio cDNA clone ZRPF clone CHBOP576F21214Q3, mRNA sequence.
ACCESSION
A1815377
VERSION
A1815377.1 GI:5430923
KEYWORDS
EST.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 657)
Look.A.T. and Holloway, M.
Zebrafish Kidney cDNA
Unpublished
Contact: Thomas Look
Dana-Farber Cancer Institute
Pediatric Oncology Department, Mayer Building-630, 44 Binney Street
, Boston, MA 02115, USA
Tel: (617) 632-5826

```

```

Fax: (617) 632-6989
Email: thomas.look@dfci.harvard.edu
The clone was obtained from an RZPD array made from an adult kidney
cDNA random primed library from Leonard Zon. RZPD library number
576. This clone is available from the Resource Centre/Primary
Database of the German Human Genome Project: RZPD (Resource
Zentrum Primär Datenbank). http://www.rzpd.de
Seq primer: T7.
FEATURES
Location/Qualifiers
source
1..657
/organism="Danio rerio"
/mol_type="rRNA"
/db_xref="taxon:7955"
/clone="RZPD clone CHBOP576F21214Q3"
/dev_stage="adult"
/clone_lib="Zebrafish Kidney cDNA random primed, RZPD
library no: 576"
/note="Organ: kidney; Vector: pBK-CMV"
BASE COUNT
163 a 165 c 162 g 165 t 2 others
ORIGIN
Query Match 25.4%; Score 127; DB 9; Length 657;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 589 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 530
Qy 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 470
Qy 137 CACCCCA 143
Db 469 CACCCCA 463
RESULT 15
AV681509/c
LOCUS
DEFINITION
661 bp mRNA linear EST 16-JAN-2002
AV681509 GKA Homo sapiens cDNA clone GKAAAC02 5', mRNA sequence.
ACCESSION
AV681509
VERSION
AV681509.1 GI:10283372
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 661)
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE
21625106
PUBMED
11752456
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
source
1..661
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"

```

```
/clone="GKAAAC02"  
/tissue type="hepatocellular carcinoma"  
/dev stage="Adult"  
/lab_host="SOLR"  
/clone_lib="GKA"  
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:  
XhoI"
```

BASE COUNT 160 a 170 c 166 g 165 t

ORIGIN

```
Query Match 25.4%; Score 127; DB 9; Length 661;  
Best Local Similarity 100.0%; Pred. No. 6.9e-58;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76  
DB 422 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 363  
  
QY 77 TTTTGGCAGTACATCAATGGGGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136  
DB 362 TTTTGGCAGTACATCAATGGGGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 303  
  
QY 137 CACCCCA 143  
DB 302 CACCCCA 296
```

Search completed: February 16, 2004, 03:15:41
Job time : 1319.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:10:26 ; Search time 3961.99 Seconds
(without alignments)
10335.846 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000

Perfect score: 1001

Sequence: 1 tctcaggtaccagtgagg.....ttcggtgcggcagcggtta 1001

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.ov.*
22: em.pat.*
23: em.ph.*
24: em.pl.*
25: em.ro.*
26: em.sts.*
27: em.un.*
28: em.vi.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_nam.*
37: em.htg_vit.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	679	67.8	278936	2	AC098563	AC098563 Rattus no
2	679	67.8	288862	2	AC099449	AC099449 Rattus no
3	494	49.4	1146	10	RNINS2	V01243 Rat Gene fo
4	358	35.8	3979	6	A30511	A30511 Plasmid p26
5	358	35.8	3980	6	A30509	A30509 Plasmid p26
6	343	34.3	2852	10	RATINSII	J00748 Rat insulin
7	291	29.1	9632	6	BD085143	BD085143 Ricin-lik
8	290	29.0	450	6	E33315	E33315 Compound al
9	290	29.0	584	6	AR272910	AR272910 Sequence
10	290	29.0	584	6	AR276491	AR276491 Sequence
11	290	29.0	584	6	AX063012	AX063012 Sequence
12	290	29.0	584	6	AX367929	AX367929 Sequence
13	290	29.0	655	6	AX351379	AX351379 Sequence
14	290	29.0	701	6	AR072501	AR072501 Sequence
15	290	29.0	701	6	AR274302	AR274302 Sequence
16	290	29.0	713	6	AR072503	AR072503 Sequence
17	290	29.0	713	6	AR274304	AR274304 Sequence
18	290	29.0	716	6	A85396	A85396 Sequence 55
19	290	29.0	716	6	AR154889	AR154889 Sequence
20	290	29.0	716	6	E65414	E65414 Genome DNA
21	290	29.0	717	6	AR072502	AR072502 Sequence
22	290	29.0	734	4	AF261959	AF261959 Anourosor
23	290	29.0	734	4	AF261959	AF261959 Anourosor
24	290	29.0	738	6	AR036905	AR036905 Sequence
25	290	29.0	738	6	AR141144	AR141144 Sequence
26	290	29.0	738	6	AR181919	AR181919 Sequence
27	290	29.0	739	4	AF261960	AF261960 Anourosor
28	290	29.0	742	6	AX197652	AX197652 Sequence
29	290	29.0	759	8	BN489202	BN489202 Beta nana
30	290	29.0	768	4	AF261962	AF261962 Anourosor
31	290	29.0	770	6	BD124786	BD124786 Primer fo
32	290	29.0	770	6	BD126222	BD126222 Primer fo
33	290	29.0	788	4	AF261963	AF261963 Anourosor
34	290	29.0	803	6	AX283744	AX283744 Sequence
35	290	29.0	809	6	AR244231	AR244231 Sequence
36	290	29.0	834	8	BN489201	BN489201 Beta nana
37	290	29.0	848	6	AR244167	AR244167 Sequence
38	290	29.0	910	9	LDO244003	LDO244003 Lepilemur
39	290	29.0	972	6	AX046223	AX046223 Sequence
40	290	29.0	1024	6	AR244042	AR244042 Sequence
41	290	29.0	1738	6	AR022240	AR022240 Sequence
42	290	29.0	1738	6	AR031416	AR031416 Sequence
43	290	29.0	1738	6	AR072989	AR072989 Sequence
44	290	29.0	1738	6	AR078254	AR078254 Sequence
45	290	29.0	1738	6	AR154630	AR154630 Sequence

ALIGNMENTS

RESULT 1
AC098563
LOCUS AC098563
DEFINITION Rattus norvegicus clone CH230-123A15, WORKING DRAFT SEQUENCE.
ACCESSION AC098563
VERSION AC098563.5 GI:30581471
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 278936)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsebrooks,S., Amin,A., Anguiano,D.,

AC098563 278936 bp DNA linear HTG 13-MAY-2003
Rattus norvegicus clone CH230-123A15, WORKING DRAFT SEQUENCE.

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, J., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, W.L., Davis, C., Dwy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, P., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hui, Y., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseg, H., Lozano, R., Lu, X., Ma, J., Mangum, B., Mapua, P., Martin, K., Mahmoud, M., Malloy, K., Mangum, A., Maheshwari, M., Mahindartne, M., Martin, R., Martinez, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwale, O., Okwunnu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Primus, E., Pu, L.-L., Plopper, F., Poldexter, A., Popovic, D., Reeves, K., Regier, M.A., Reigh, R., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vallas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Submitted (24-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 278936)

AUTHORS

TITLE

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center	
Center: Baylor College of Medicine	
Center code: BCM	
Web site: http://www.hgsc.bcm.tmc.edu/	
Contact: hgsc-help@bcm.tmc.edu	
----- Project Information	
Center project name: GIKN	
Center clone name: CH230-123A15	
----- Summary Statistics	
Assembly program: Atlas 3.0.	
Consensus quality: 257140 bases at least Q40	
Consensus quality: 257743 bases at least Q30	
Consensus quality: 258080 bases at least Q20	
Estimated insert size: 269272; sum-of-contigs estimation	
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation	

* NOTE: Estimated insert size may differ from sequence length	
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)	
* NOTE: This sequence may represent more than one clone.	
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 1 contigs. Gaps between the contigs	
* are represented as runs of N. The order of the pieces	
* is believed to be correct as given, however the sizes	
* of the gaps between them are based on estimates that have	
* provided by the submitter.	
* This sequence will be replaced	
* by the finished sequence as soon as it is available and	
* the accession number will be preserved.	
* 1 278936: contig of 278936 bp in length.	

FEATURES	
source	
1..278936	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clones="CH230-123A15"	
1..1392	
misc_feature	
/note="wgs end extension	
clone_end:Sp6"	
7896..8641	
/note="clone boundary	
clone_end:Sp6	
site:EcoRI	
end_sequence:BH294789"	
276008..277823	
/note="wgs contig"	
277874..278936	
/note="wgs contig"	
BASE COUNT 65595 a 64123 c 64880 g 63888 t 20450 others	
ORIGIN	
Query Match 67.8%; Score 679; DB 2; Length 278936;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 TCTAGGCTTAACCATGGGGGCGACATTTCTGTGGGAGCTAGACATATGTAACATGGT 60	
Db 47355 TCTAGGCTTAACCATGGGGGCGACATTTCTGTGGGAGCTAGACATATGTAACATGGT 47414	
QY 61 AGCTGCCAGGAGGAGTGGAGTCTCTCTTAAGTCTCTAGTGGTGGCTAG 120	
Db 47415 AGCTGCCAGGAGGAGTGGAGTCTCTCTTAAGTCTCTAGTGGTGGCTAG 47474	
QY 121 GCCCAGGATAGGTACCTATTATTTGGGGAGCCCATAGACACTGCACTGATGAGGATGGT 180	
Db 47475 GCCCAGGATAGGTACCTATTATTTGGGGAGCCCATAGACACTGCACTGATGAGGATGGT 47534	
QY 181 AACAGATGTGTAGTTTGGAGGCCCATATGCTCCATTCATCAGCTGACTTGTCTCAC 240	
Db 47535 AACAGATGTGTAGTTTGGAGGCCCATATGCTCCATTCATCAGCTGACTTGTCTCAC 47594	
QY 241 AGCCATGCAACCTTGCCTCTCTGCTGCTGCTAGCTTTAGCAGGGGATAAAGTGAGAGAAACCTGT 300	

* as soon as it is available and the accession number will
* be preserved.

1 281836: contig of 281836 bp in length
281837 281936: gap of unknown length
281937 282992: contig of 1056 bp in length
282993 283092: gap of unknown length
283093 284135: contig of 1043 bp in length
284136 284236: gap of unknown length
284237 285379: contig of 1144 bp in length
285380 285479: gap of unknown length
285480 286963: contig of 1484 bp in length
286964 287063: gap of unknown length
287064 288862: contig of 1799 bp in length.

FEATURES

source

Location/Qualifiers
1..288862
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-57B2"
1..2458
/note="wgs contig"
complement(280810..281268)
/note="clone boundary
clone end:Sp5
site:ECORI
end sequence:BH296258"

BASE COUNT 75546 a 55666 c 58691 g 65812 t 33147 others
ORIGIN

Query Match 67.8%; Score 679; DB 2; Length 288862;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 60
Db 226123 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 226182

Qy 61 AGCTGCCAGGAAGAGTGAATCTTCTTAAAGTCTCTAGTGTGACGGGTGGCTAG 120
Db 226183 AGCTGCCAGGAAGAGTGAATCTTCTTAAAGTCTCTAGTGTGACGGGTGGCTAG 226242

Qy 121 GCCCCAGGATAGTACTATTGGGACCCCATAGAGCACTGCACCTAGCTAGGGATGT 180
Db 226243 GCCCCAGGATAGTACTATTGGGACCCCATAGAGCACTGCACCTAGCTAGGGATGT 226302

Qy 181 AACAGGATGTAGTTTGGAGCCCATATGCCATTATGACCACTGACTTGTCTCAC 240
Db 226303 AACAGGATGTAGTTTGGAGCCCATATGCCATTATGACCACTGACTTGTCTCAC 226362

Qy 241 AGCCATGCAACCCCTGCTCTCTGCTGCTGACTTAGCAGGGGATAAAGTGAGAGAAAGCCTG 300
Db 226363 AGCCATGCAACCCCTGCTCTCTGCTGCTGACTTAGCAGGGGATAAAGTGAGAGAAAGCCTG 226422

Qy 301 GGCTAATCAGGGGTGCTCAGCTCCTCCTAACTGATTTGCTATGTTCTTTGCTTCT 360
Db 226423 GGCTAATCAGGGGTGCTCAGCTCCTCCTAACTGATTTGCTATGTTCTTTGCTTCT 226482

Qy 361 GTGCTGTGATGCTCTGCCCTGTGCTGACATGACCTCCCTGGCAGTGCGCACTGGAGC 420
Db 226483 GTGCTGTGATGCTCTGCCCTGTGCTGACATGACCTCCCTGGCAGTGCGCACTGGAGC 226542

Qy 421 TGGGTGAGGCGCGGGGAGGTGACCTTCAGACCTTGGCACTGGAGGTGGCCCGGACAG 480
Db 226543 TGGGTGAGGCGCGGGGAGGTGACCTTCAGACCTTGGCACTGGAGGTGGCCCGGACAG 226602

Qy 481 AGCCGCGCATCTGGATCAGTGTGTCACCAAGCATCTCTCTCTCTCACTCACTGGAACT 540
Db 226603 AGCCGCGCATCTGGATCAGTGTGTCACCAAGCATCTCTCTCTCTCACTCACTGGAACT 226662

Qy 541 ACTGCACTAGGCGCCACCATACCTCTGTCACCCCTCTGCAATGAATAAAACCTTTGAAA 600
Db 226663 ACTGCACTAGGCGCCACCATACCTCTGTCACCCCTCTGCAATGAATAAAACCTTTGAAA 226722

Qy 601 GAGCACTACAAAGTTGTGTACATGCGTGCATGTGCATATGTGTGTGGGGGAAACATGA 660
Db 226723 GAGCACTACAAAGTTGTGTACATGCGTGCATGTGCATATGTGTGTGGGGGAAACATGA 226782

Qy 661 GTGGGGCTGGCTGGAGTGG 679
Db 226783 GTGGGGCTGGCTGGAGTGG 226801

RESULT 3

RNINS2

LOCUS

DEFINITION

Rat gene for insulin 2.

ACCESSION

V01243

VERSION

V01243.1

KEYWORDS

insulin; preproinsulin.

SOURCE

Rattus norvegicus

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 1146)

AUTHORS

Lomedico P., Rosenthal N., Efstratidis A., Gilbert W.,

Kolodner R. and Tizard R.

The structure and evolution of the two nonallelic rat preproinsulin

Genes

Cell 18 (2), 545-558 (1979)

MEDLINE

80045035

PUBMED

498284

REFERENCE

2

AUTHORS

Kozak M.

Point mutations close to the AUG initiator codon affect the

efficiency of translation of rat preproinsulin in vivo

Nature 308 (5956), 241-246 (1984)

MEDLINE

84142290

PUBMED

6700727

FEATURES

Location/Qualifiers

1..1146

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

prim_transcript

43..1103

exon

43..82

/note="messenger RNA"

/number=1

83..201

/number=1

202..404

/number=2

join(219..404,904..1050)

/codon_start=1

/product="preproinsulin"

/protein_id="CAA24560.1"

/db_xref="GI:56490"

/db_xref="SWISS-PROT:P01323"

/translation="MALWIFLPLALLIWEPRPAQAFVKHLCQSHLVEALYLVCGL

ERGFYTPMGRREVDEDPQAQLGLGGPGAGDQLTLALEVARQKRGIVDQCCTSI

YQLENYCN"

intron

405..903

/number=2

904..1103

/number=3

BASE COUNT

236 a 313 c 308 g 289 t

ORIGIN

Query Match

Best Local Similarity

Matches 644; Conservative

49.4%; Score 494; DB 10; Length 1146;

Pred. No. 6.4e-278;

Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 60

Db 500 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 559


```
QY 61 AGCTCCAGGAAGAGTGAATCTCTTAAAGTCTCTAGTGGTGACGGGTGGCTAG 120
Db 560 AGCTGCCAAGAGGAGTGAATCTCTTAAAGTCTCTAGTGGTGACGGGTGGCTAG 619
QY 121 GCCCAGGATAGTACCTATTTGGGAGCCCAATAGAGCACTGCATGACTGAGGATGGT 180
Db 620 GCCCAGGATAGTACCTATTTGGGAGCCCAATAGAGCACTGCATGACTGAGGATGGT 679
QY 181 AACAGGATGTAGTATTTGGAGGCCCATATATCCATTATGATGACAGGATGCTGCTCAC 240
Db 680 AACAGGATGTAGTATTTGGAGGCCCATATATGCTCATGACAGGATGCTGCTCAC 739
QY 241 AGCCATGCAACCTTGGCTCTCTGCTGACCTTAGCAGGAGTAAAGTACAGAGAGCCCTG 300
Db 740 AGCCATGCAACCTTGGCTCTCTGCTGACCTTAGCAGGAGTAAAGTACAGAGAGCCCTG 799
QY 301 GGTCTAATCAGGGGGTGGCTCAGCTCTCTTAACTGATGCTCTTAACTGATGCTTCTTCT 360
Db 800 GGTCTAATCAGGGGGTGGCTCAGCTCTCTTAACTGATGCTCTTAACTGATGCTTCTTCT 859
QY 361 GTGCTGCTGATGCTTGGCTCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 420
Db 860 GTGCTGCTGATGCTTGGCTCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 919
QY 421 TGCGTGGAGCCCGGGGGCGGTGACCTTTCAGACCTTGGCACTGAGGTGGCCGGCAGA 480
Db 920 TGCGTGGAGCCCGGGGGCGGTGACCTTTCAGACCTTGGCACTGAGGTGGCCGGCAGA 979
QY 481 AGCGGGGATCTGGATCAGTCTGCTGACCAAGCATCTGCTCTCTTACCACTGGAGAACT 540
Db 980 AGCGGGGATCTGGATCAGTCTGCTGACCAAGCATCTGCTCTCTTACCACTGGAGAACT 1039
QY 541 ACTGCACTAGGCCCACTACCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 600
Db 1040 ACTGCACTAGGCCCACTACCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1099
QY 601 GAGCACTACAAGTGTGTATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCT 647
Db 1100 GAGCACTACAAGTGTGTATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1146

RESULT 4
A30511
LOCUS Plasmid p264BAM. 3979 bp DNA linear PAT 05-JUL-2002
DEFINITION A30511
ACCESSION A30511
VERSION A30511.1 GI:23957137
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM F. Hoffmann-La Roche AG
REFERENCE 1 (bases 1 to 3979)
AUTHORS Fountoulakis, M., Garotta, G. and Stueber, D.
TITLE Soluble interferon-gamma receptors and methods for their production
JOURNAL Patent: EP 039502-A 78 24-OCT-1990;
F. Hoffmann-La Roche AG
FEATURES
source
1..3979
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 946 a 988 c 1115 g 930 t
ORIGIN
Query Match 35.8%; Score 358; DB 6; Length 3979;
Best Local Similarity 99.6%; Pred. No. 5e-198;
Matches 458; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 171 GAGGATGTGTACAGGATGTAGTATTTGGAGGCCCATATGTCATGACCAAGTGA 230
Db 873 GAGGATGTGTACAGGATGTAGTATTTGGAGGCCCATATGTCATGACCAAGTGA 932
QY 231 CTTGTCTCAGAGCATGCAACCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 290
Db 994 AGAAGCCTGGGCTAATCGGGGGTGCCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1053
QY 351 CTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
Db 1054 CTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113
QY 411 CAACTGAGCTGGGTGGAGCCCGGGGGAGGTGACCTTTCAGACCTTGGCACTGGAGGTG 470
```

```
Db 933 CTTGTCTCAGAGCATGCAACCTTGCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 992
QY 291 AGAAGCCTGGGCTAATCAGGAGGTGCTCAGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 350
Db 993 AGAAGCCTGGGCTAATCAGGAGGTGCTCAGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 1052
QY 351 CTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
Db 1053 CTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1112
QY 411 CAACTGAGCTGGGTGGAGCCCGGGGGAGGTGACCTTTCAGACCTTGGCACTGGAGGTG 470
Db 1113 CAACTGAGCTGGGTGGAGCCCGGGGGAGGTGACCTTTCAGACCTTGGCACTGGAGGTG 1172
QY 471 GCCCGGCAGAGCGCGGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 530
Db 1173 GCCCGGCAGAGCGCGGCATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1232
QY 531 CTGGAGAACTACTGCAACTAGGCCCACTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
Db 1233 CTGGAGAACTACTGCAACTAGGCCCACTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
QY 591 ACCTTTGAAAGAGCACTACAAAGTTGTGTATCATGCGTGC 630
Db 1293 ACCTTTGAAAGAGCACTACAAAGTTGTGTATCATGCGTGC 1332

RESULT 5
A30509
LOCUS Plasmid p264BGL. 3980 bp DNA linear PAT 05-JUL-2002
DEFINITION A30509
ACCESSION A30509
VERSION A30509.1 GI:23957136
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM F. Hoffmann-La Roche AG
REFERENCE 1 (bases 1 to 3980)
AUTHORS Fountoulakis, M., Garotta, G. and Stueber, D.
TITLE Soluble interferon-gamma receptors and methods for their production
JOURNAL Patent: EP 039502-A 78 24-OCT-1990;
F. Hoffmann-La Roche AG
FEATURES
source
1..3980
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 947 a 987 c 1114 g 931 t 1 others
ORIGIN
Query Match 35.8%; Score 358; DB 6; Length 3980;
Best Local Similarity 99.6%; Pred. No. 5e-198;
Matches 458; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 171 GAGGATGTGTACAGGATGTAGTATTTGGAGGCCCATATGTCATGACCAAGTGA 230
Db 874 GAGGATGTGTACAGGATGTAGTATTTGGAGGCCCATATGTCATGACCAAGTGA 933
QY 231 CTTGTCTCAGAGCATGCAACCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 290
Db 934 CTTGTCTCAGAGCATGCAACCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 993
QY 291 AGAAGCCTGGGCTAATCAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 350
Db 994 AGAAGCCTGGGCTAATCGGGGGTGCCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1053
QY 351 CTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
Db 1054 CTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113
QY 411 CAACTGAGCTGGGTGGAGCCCGGGGGAGGTGACCTTTCAGACCTTGGCACTGGAGGTG 470
```



```

RESULT 7
BD085143
LOCUS          9632 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION    Ricin-like toxin variants for treatment of cancer, viral or
               parasitic infections.
ACCESSION     BD085143
VERSION       BD085143.1 GI:22630753
KEYWORDS      JP 2001523961-A/1.
SOURCE        unidentified
ORGANISM      unidentified
               unclassified.
REFERENCE     1 (bases 1 to 9632)
AUTHORS      Borgford,T.
TITLE        Ricin-like toxin variants for treatment of cancer, viral or
               parasitic infections
JOURNAL       TWISTLAND THERAPEUTICS INC
COMMENT      OS Baculovirus transfer vector pVL 1393
               PN JP 2001523961-A/1
               PD 27-NOV-2001
               PF 30-APR-1998 JP 1998546437
               PR 30-APR-1997 US 60/045148,29-OCT-1997 US 60/063715 PI
               PC C12N15/29,C12N15/62,C12N15/70,C12N15/86,A61K48/00 CC
               Ricin-like toxin variants for treatment of cancer, viral or CC
               parasitic
               CC infections
               FH Key
               FT source
               FT 1.9632
               Location/Qualifiers
               1.9632
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"
BASE COUNT    2602 a 2122 c 2176 g 2732 t
ORIGIN
Query Match 29.1%; Score 291; DB 6; Length 9632;
Best Local Similarity 100.0%; Pred. No. 1.1e-158;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 711 ACCTAATCATGTCTAGCTGTTTCTGTGTGAATTTGTTATTCGCTCACAATTCACAC 770
Db 7002 ACCTAATCATGTCTAGCTGTTTCTGTGTGAATTTGTTATTCGCTCACAATTCACAC 7061
QY 771 AACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGTAACTC 830
Db 7062 AACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGTAACTC 7121
QY 831 ACATTAATTCGTTGCGTCACTGCCCGCTTTCAGTCGGGAAACCTGTGTCGCGAGCTG 890
Db 7122 ACATTAATTCGTTGCGTCACTGCCCGCTTTCAGTCGGGAAACCTGTGTCGCGAGCTG 7181
QY 891 CATTAATGAATCGGCAACGCGCGGAGAGCGGTTTTCGCTATTGGGCGCTCTTCGCT 950
Db 7182 CATTAATGAATCGGCAACGCGCGGAGAGCGGTTTTCGCTATTGGGCGCTCTTCGCT 7241
QY 951 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
Db 7242 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 7292

RESULT 8
E33315
LOCUS          450 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION    Compound alkylating specific base sequence in DNA and process for
               synthesizing the same.
ACCESSION     E33315
VERSION       E33315.1 GI:18624059
KEYWORDS      JP 2000159768-A/2.
SOURCE        unidentified

```

```

ORGANISM      unidentified
               unclassified.
REFERENCE     1 (bases 1 to 450)
AUTHORS      Sugiyama,H., To,S. and Saito,A.
TITLE        Compound alkylating specific base sequence in DNA and process for
               synthesizing the same
JOURNAL       SCIENCE & TECH AGENCY
COMMENT      OS EUC 18
               PN JP 2000159768-A/2
               PD 13-JUN-2000
               PF 14-SEP-1998 JP 1998260710
               PR HIROSHI SUGIYAMA,SHIFUKU TO,AKIRA SAITO
               PC C07D487/04,A61K48/00,C12N15/09,C12N15/00
               CC CC
               FH Key
               FT source
               FT 1.450
               Location/Qualifiers
               1.450
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"
BASE COUNT    107 a 119 c 124 g 100 t
ORIGIN
Query Match 29.0%; Score 290; DB 6; Length 450;
Best Local Similarity 100.0%; Pred. No. 4.6e-158;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CQTAAATCATGTCTAGCTGTTTCTGTGTGAATTTGTTATTCGCTCACAATTCACACA 771
Db 78 CQTAAATCATGTCTAGCTGTTTCTGTGTGAATTTGTTATTCGCTCACAATTCACACA 137
QY 772 ACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGTAACTCA 831
Db 138 ACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGTAACTCA 197
QY 832 CATTAATTCGTTGCGTCACTGCCCGCTTTCAGTCGGGAAACCTGTGTCGCGAGCTGC 891
Db 198 CATTAATTCGTTGCGTCACTGCCCGCTTTCAGTCGGGAAACCTGTGTCGCGAGCTGC 257
QY 892 ATTAATGAATCGGCAACGCGCGGAGAGCGGTTTTCGCTATTGGGCGCTCTTCGCTT 951
Db 258 ATTAATGAATCGGCAACGCGCGGAGAGCGGTTTTCGCTATTGGGCGCTCTTCGCTT 317
QY 952 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
Db 318 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 367

RESULT 9
AR272910
LOCUS          584 bp      DNA      linear      PAT 10-APR-2003
DEFINITION    Sequence 639 from patent US 6504010.
ACCESSION     AR272910
VERSION       AR272910.1 GI:29704795
KEYWORDS
SOURCE        Unknown.
ORGANISM      Unknown.
               unclassified.
REFERENCE     1 (bases 1 to 584)
AUTHORS      Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
               Carter,D., Retter,M.W., Mannion,J. and Fan,L.
TITLE        Compositions and methods for the therapy and diagnosis of lung
               cancer
JOURNAL       Patent: US 6504010-A 639 07-JAN-2003;
COMMENT      Location/Qualifiers
               1.584
               /organism="unknown"
BASE COUNT    142 a 159 c 153 g 130 t
ORIGIN

```



```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 142 a 159 c 153 g 130 t
ORIGIN
Query Match 29.0%; Score 290; DB 6; Length 584;
Best Local Similarity 100.0%; Pred. No. 4.6e-158;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTAAGTGTGAAATGTTATCCGCTCAAAATCCACACA 771
Db 37 CGTAATCATGTCATAGCTGTTTCCTGTGTAAGTGTGAAATGTTATCCGCTCAAAATCCACACA 96
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGCGGCTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGCGGCTGCTTAATGAGTGAGCTAACTCA 156
QY 832 CATTAAATGCTGCTCACTCCCGCTTTCAGTCGGGAACCTGTCGCGAGCTGC 891
Db 157 CATTAAATGCTGCTCACTCCCGCTTTCAGTCGGGAACCTGTCGCGAGCTGC 216
QY 892 ATTAATGAATCGCCAAACGCGGAGAGCGGTTTGGTATTTGGCTGCTCTTCCGCTT 951
Db 217 ATTAATGAATCGCCAAACGCGGAGAGCGGTTTGGTATTTGGCTGCTCTTCCGCTT 276
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 326
RESULT 13
AX351379 655 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 126 from Patent WO0196390.
ACCESSION AX351379
VERSION AX351379.1 GI:18616726
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Jjiang, Y., Hepler, W.T., Clapper, J.D., Wang, A. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon
JOURNAL cancer
PATENT: WO 0196390-A 126 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 655
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 185 c 169 g 151 t
ORIGIN
Query Match 29.0%; Score 290; DB 6; Length 655;
Best Local Similarity 100.0%; Pred. No. 4.6e-158;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTAAGTGTGAAATGTTATCCGCTCAAAATCCACACA 771
Db 31 CGTAATCATGTCATAGCTGTTTCCTGTGTAAGTGTGAAATGTTATCCGCTCAAAATCCACACA 90
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGCGGCTGCTTAATGAGTGAGCTAACTCA 831
Db 91 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGCGGCTGCTTAATGAGTGAGCTAACTCA 150
QY 832 CATTAAATGCTGCTCACTCCCGCTTTCAGTCGGGAACCTGTCGCGAGCTGC 891
Db 151 CATTAAATGCTGCTCACTCCCGCTTTCAGTCGGGAACCTGTCGCGAGCTGC 210
QY 892 ATTAATGAATCGCCAAACGCGGAGAGCGGTTTGGTATTTGGCTGCTCTTCCGCTT 951
```

```
211 ATTAATGAATCGCCAAACGCGGAGAGCGGTTTGGTATTTGGCGCTCTTCCGCTT 270
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 271 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 320
RESULT 14
AR072501 701 bp DNA linear PAT 28-AUG-2000
LOCUS
DEFINITION Sequence 24 from patent US 5948614.
ACCESSION AR072501
VERSION AR072501.1 GI:9999265
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 701)
AUTHORS Chatterjee, D.K.
TITLE Cloned DNA Polymerases from thermotoga maritima and mutants thereof
JOURNAL Patent: US 5948614-A 24 07-SEP-1999;
FEATURES
source
1. 701
/organism="unknown"
BASE COUNT 154 a 179 c 191 g 152 t 25 others
ORIGIN
Query Match 29.0%; Score 290; DB 6; Length 701;
Best Local Similarity 100.0%; Pred. No. 4.6e-158;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTAAGTGTGAAATGTTATCCGCTCAAAATCCACACA 771
Db 47 CGTAATCATGTCATAGCTGTTTCCTGTGTAAGTGTGAAATGTTATCCGCTCAAAATCCACACA 106
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGCGGCTGCTTAATGAGTGAGCTAACTCA 831
Db 107 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGCGGCTGCTTAATGAGTGAGCTAACTCA 166
QY 832 CATTAAATGCTGCTCGCTCACTCCCGCTTTCAGTCGGGAACCTGTCGCGAGCTGC 891
Db 167 CATTAAATGCTGCTCGCTCACTCCCGCTTTCAGTCGGGAACCTGTCGCGAGCTGC 226
QY 892 ATTAATGAATCGCCAAACGCGGAGAGCGGTTTGGTATTTGGCTGCTCTTCCGCTT 951
Db 227 ATTAATGAATCGCCAAACGCGGAGAGCGGTTTGGTATTTGGCTGCTCTTCCGCTT 286
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 287 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 336
RESULT 15
AR274302 701 bp DNA linear PAT 10-APR-2003
LOCUS
DEFINITION Sequence 24 from patent US 6506560.
ACCESSION AR274302
VERSION AR274302.1 GI:29706732
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 701)
AUTHORS Hughes, A.J. and Chatterjee, D.K.
TITLE Cloned DNA polymerases from Thermotoga and mutants thereof
JOURNAL Patent: US 6506560-A 24 14-JAN-2003;
FEATURES
source
1. 701
/organism="unknown"
BASE COUNT 154 a 179 c 191 g 152 t 25 others
ORIGIN
Query Match 29.0%; Score 290; DB 6; Length 701;
```

Best Local Similarity 100.0%; Pred. No. 4.6e-158;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	712	CGTAATCATGGTCAATAGCTGTTTCCTGTGTGAAATTTGTATCCGCTCACAATTCACACA	771
Db	47	CGTAATCATGGTCAATAGCTGTTTCCTGTGTGAAATTTGTATCCGCTCACAATTCACACA	106
QY	772	ACATACGAGCCGAGAGCATTAAGCTGAAGCTGGGTGCTATAGTGAAGTCACTCA	831
Db	107	ACATACGAGCCGAGAGCATTAAGCTGAAGCTGGGTGCTATAGTGAAGTCACTCA	166
QY	832	CATTAAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	891
Db	167	CATTAAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	226
QY	892	ATTATGATTCGGCCCAACGCGGGGAGAGCGGTTTTCGTATTGGGGCTCTTCGCTT	951
Db	227	ATTATGATTCGGCCCAACGCGGGGAGAGCGGTTTTCGTATTGGGGCTCTTCGCTT	286
QY	952	CCTCGCTCACTGACTGCTGCTCGCTCGGTGCTGCGCTGCGCGAGCGGTA	1001
Db	287	CCTCGCTCACTGACTGCTGCTGCTCGGTGCTGCGCTGCGCGAGCGGTA	336

Search completed: February 16, 2004, 01:25:51
Job time : 3967.99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:07:11 ; Search time 324.065 Seconds
(without alignments)
8338.264 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000
Perfect score: 1001
Sequence: 1 tctcagtgtaaccagtgagg.....ttcggctgcggcagcggtta 1001

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: listing first 45 summaries

- Database : N_Geneseq_19Jun03.*
- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
 - 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
 - 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
 - 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
 - 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
 - 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
 - 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
 - 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
 - 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
 - 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
 - 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
 - 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
 - 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
 - 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
 - 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
 - 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
 - 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
 - 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
 - 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
 - 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
 - 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
 - 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
 - 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
 - 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
 - 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	5283	ABK10062	Expression vector
2	309	30.9	3979	AA006312	Sequence of plasmid
3	307	30.7	5552	AA58309	pDIP/PyCSP.1 plasm
4	291	29.1	9632	AA979598	Baculovirus transf
5	291	29.1	9632	AA04251	Baculovirus transf
6	290	29.0	409	ABV45293	Human prostate exp
7	290	29.0	450	AAA10990	Fragment #2 of pUC
8	290	29.0	515	AAA82004	N. meningitidis pa

9	290	29.0	561	23	ABV45308	Human prostate exp
10	290	29.0	584	22	AAF68704	Human lung tumour
11	290	29.0	584	24	ABK38615	cDNA encoding clon
12	290	29.0	584	25	ACA10944	Human lung neuroen
13	290	29.0	584	25	ABX99895	Lung cancer therap
14	290	29.0	598	21	AAK62566	Human breast cance
15	290	29.0	598	24	ABT12681	Orestes sequencing
16	290	29.0	655	24	AKZ27689	Human colon cancer
17	290	29.0	686	22	ABA20358	Human nervous syst
18	290	29.0	686	22	ABA21448	Human nervous syst
19	290	29.0	686	22	AA105999	Human reproductive
20	290	29.0	686	22	AA107025	Human reproductive
21	290	29.0	686	22	AA527722	DNA encoding novel
22	290	29.0	686	22	AAK81476	Human immune/haema
23	290	29.0	686	22	AA533387	DNA encoding human
24	290	29.0	686	23	ABL98563	Human testicular a
25	290	29.0	686	24	ABK69910	Human secreted pro
26	290	29.0	686	25	ABZ74009	Secreted protein g
27	290	29.0	686	25	ABZ74600	Secreted protein g
28	290	29.0	686	25	ABZ68007	Human secreted pro
29	290	29.0	686	25	ABZ68121	Human secreted pro
30	290	29.0	691	21	AAA81954	N. meningitidis pa
31	290	29.0	717	21	AAZ88769	T. neapolitana Tne
32	290	29.0	721	21	AAA81967	N. meningitidis pa
33	290	29.0	728	21	AAA81752	N. meningitidis pa
34	290	29.0	738	20	AAZ09716	Plasmid pGEM DNA I
35	290	29.0	739	22	AAK61981	Human immune/haema
36	290	29.0	742	22	AAH82483	Human ovarian tumo
37	290	29.0	752	19	AAV31294	E. coli J96 pathog
38	290	29.0	770	22	AAK91757	Human CDNA 5'-end
39	290	29.0	770	22	AAK93193	Human CDNA clone r
40	290	29.0	803	24	AA561891	Porcine muscular s
41	290	29.0	809	21	AAZ97398	Human prostate can
42	290	29.0	848	21	AAZ97334	Human prostate can
43	290	29.0	972	21	AAK83524	T7 promoter sequen
44	290	29.0	1024	21	AAZ97209	Human prostate can
45	290	29.0	1091	23	ABK43594	DNA encoding novel

ALIGNMENTS

RESULT 1
ABK10062
ID ABK10062 standard; DNA; 5283 BP.

XX AC ABK10062;
XX DT 21-MAY-2002 (first entry)
XX DE Expression vector construct pVGI.1 containing VEGF-2 insert.

XX KW Human; vascular endothelial growth factor 2; VEGF-2; ds; pVGI.1;
XX KW chronic limb ischaemia; myocardial ischaemia; autoimmune disorder; pHE4a;
XX KW allergic reaction; organ rejection; inflammatory condition; arrhythmia;
XX KW hyperproliferative disorder; viral infection; bacterial infection;
XX KW fungal infection; parasitic infection; cardiovascular disorder; embolism;
XX KW heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.

XX OS Synthetic.
XX PN WO200211769-A1.

XX PD 14-FEB-2002.
XX PF 03-AUG-2001; 2001WO-US24658.
XX PR 04-AUG-2000; 2000US-223276P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Coleman TA;
XX

PA (USNA) US SEC OF NAVY.

KW Ricin; toxin; antiviral; virucide; retrovirus; protease; HIV-I;

KW cancer; infection; therapy; vector; pVL1393; ds; circular; cyclic.
 XX. Baculovirus.
 OS WC9741233-A1.
 XX 06-NOV-1997.
 XX 29-APR-1997; 97W0-CA00288.
 XX 30-APR-1996; 96US-0016509.
 XX (CANG-) CANGENE CORP.
 XX Borgford T;
 XX WPI; 1997-549735/50.
 XX DNAs encoding ricin like toxins A and B - are linked via linker
 PT containing cleavage site for retroviral protease, used to inhibit or
 PT destroy mammalian cells infected with retrovirus
 XX
 XX Claim 13; Fig 11; 105pp; English.
 XX This is the DNA sequence of baculovirus transfer vector pVL1393.
 CC A claimed baculovirus transfer vector incorporates a nucleic acid
 CC (see AAR97899-901, AAR97910, AAR97913, AAR97916 and AAR97919) that
 CC encodes a recombinant protein in which the A and B chains of a
 CC ricin-like toxin are linked by a heterologous linker peptide (see
 CC AAM3680-86) that contains a cleavage recognition site for a retroviral
 CC protease such as HIV, HTLV-I or HTLV-II. The vector provides expression
 CC of the recombinant proteins in Sf9 insect cells. The recombinant
 CC proteins can be used to selectively inhibit or destroy mammalian
 CC cells infected with a retrovirus, such as cancer cells associated
 CC with HTLV or cells associated with HIV. The recombinant proteins
 CC are non-toxic until the ricin A chain is liberated from the B chain
 CC by the retroviral protease, and thus can be used to specifically
 CC target infected cells without the need for a cell binding
 CC component.
 XX
 XX Sequence 9632 BP; 2602 A; 2122 C; 2176 G; 2732 T; 0 other;
 Query Match 29.1%; Score 291; DB 18; Length 9632;
 Best Local Similarity 100.0%; Pred. NO. 2e-81;
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 711 ACCTAATCATGTCATAGCTGTTTCTGTCGAAATTTGTTATCCGCTCACAATTCACAC 770
 DB 4302 ACCTAATCATGTCATAGCTGTTTCTGTCGAAATTTGTTATCCGCTCACAATTCACAC 4361
 QY 771 AACATACGCGCGAGCATTAAGTCTAAGCTCGGCTGCTATGATGAGTCACTC 830
 DB 4362 AACATACGCGCGAGCATTAAGTCTAAGCTCGGCTGCTATGATGAGTCACTC 4421
 QY 831 ACATTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTG 890
 DB 4422 ACATTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTG 4481
 QY 891 CATTAATGATCGGCAACGCGCGGAGAGCGGTTTCGTTATGCGGCTCTTCGCT 950
 DB 4482 CATTAATGATCGGCAACGCGCGGAGAGCGGTTTCGTTATGCGGCTCTTCGCT 4541
 QY 951 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
 DB 4542 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4592
 RESULT 5
 AAX04251
 ID AAX04251 standard; DNA; 9632 BP.
 XX
 XX AAX04251;
 XX

DT 15-APR-1999 (first entry)
 XX Baculovirus transfer vector pVL1393 DNA sequence.
 XX Ricin-like toxin; cancer; viral infection; parasitic infection;
 KW linker; B chain; A chain; protease; fungal infection; malaria;
 XX leucocyte proliferation; cytomegalovirus; herpes; hepatitis;
 KW rhinovirus; laryngotracheitis; poliomyelitis; varicella zoster;
 KW cystic fibrosis; multiple sclerosis; ds.
 XX Escherichia coli.
 OS Synthetic.
 XX WO9849311-A2.
 XX 05-NOV-1998.
 XX 30-APR-1998; 98W0-CA00394.
 XX 29-OCT-1997; 97US-0063715.
 XX 30-APR-1997; 97US-0045148.
 XX (DNOV-) DE NOVO ENZYME CORP.
 XX Borgford T;
 XX WPI; 1999-009431/01.
 XX New nucleic acid encoding ricin-like toxin with an interchain linker
 PT cleaved by protease - is specific for diseased cells, useful for,
 PT e.g. killing selectively cancer or infected cells
 XX Disclosure; Fig 1; 352pp; English.
 XX The present invention describes new purified and isolated nucleic acids
 CC (i) encoding: (i) the A and B chains of a ricin-like toxin (ii); and
 CC (ii) a heterologous linker, joining the two chains and including a
 CC cleavage recognition site for a disease-specific protease (iii). Also
 CC described are: (1) plasmids or baculovirus transfer vectors that contain
 CC (i); and (2) recombinant protein (iv) consisting of the A and B chains
 CC of (ii) joined by the specified linker. (iv), produced by expression of
 CC (i) in host cells, are used to inhibit or kill diseased cells that
 CC produce (iii), particularly for treating cancers (e.g. leucocyte
 CC proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or
 CC infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.
 CC cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngotracheitis,
 CC poliomyelitis or varicella zoster), also cystic fibrosis and multiple
 CC sclerosis. Alternatively, (i) is used to express (iv) in vivo. (iv) is
 CC toxic specifically for (iii)-expressing cells and does not depend for
 CC specificity on a cell-binding component. When used to treat virus-
 CC infected cells, transcytosis and cytotoxicity of (iv) are increased by
 CC retrograde translocation from endoplasmic reticulum to cytoplasm (which
 CC some viruses exploit to avoid immune detection), so selectivity and
 CC safety are further improved. (iv) are not toxic until chain A is
 CC released and this occurs only in target cells. The present sequence
 CC represents a nucleotide sequence from the present invention.
 XX
 XX Sequence 9632 BP; 2602 A; 2122 C; 2176 G; 2732 T; 0 other;
 Query Match 29.1%; Score 291; DB 20; Length 9632;
 Best Local Similarity 100.0%; Pred. NO. 2e-81;
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 711 ACCTAATCATGTCATAGCTGTTTCTGTCGAAATTTGTTATCCGCTCACAATTCACAC 770
 DB 7002 ACCTAATCATGTCATAGCTGTTTCTGTCGAAATTTGTTATCCGCTCACAATTCACAC 7061
 QY 771 AACATACGCGCGAGCATTAAGTCTAAGCTCGGCTGCTATGATGAGTCACTC 830
 DB 7062 AACATACGCGCGAGCATTAAGTCTAAGCTCGGCTGCTATGATGAGTCACTC 7121
 QY 831 ACATTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTG 890

Best Local Similarity 100.0%; Pred. No. 1.2e-80; Mismatches 0; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATTCACACA 771
DB 78 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATTCACACA 137

QY 772 ACATAGAGCGCGAAGCATAAGTGTAAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 831
DB 138 ACATAGAGCGCGAAGCATAAGTGTAAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 197

QY 832 CATTAATGCGTTCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTGC 891
DB 198 CATTAATGCGTTCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTGC 257

QY 892 ATTAATGAATCGCGCAACGCGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT 951
DB 258 ATTAATGAATCGCGCAACGCGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT 317

QY 952 CTCGCTCACTAGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
DB 318 CTCGCTCACTAGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 367

RESULT 8

AA82004
ID AA82004 standard; DNA; 515 BP.
XX
AC AA82004;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_551 SEQ ID NO:551.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.

OS Neisseria meningitidis.
XX
XX
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US23573.
XX
XX 09-OCT-1998; 98US-0103794.
XX
XX 30-APR-1999; 99US-0132068.
XX
XX (CHIR) CHIRON CORP.

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Scarlato V;
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
XX WPI; 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisseria infections, for example, N.gonorrhoea
XX
XX Claim 7; Page 1653; 1760pp; English.

PS
XX The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AA81453 to AA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AA81260 to AA81303 and AA825620 to AA825663 represent
CC Neisseria DNA sequences and their corresponding proteins; AA81254 to
CC AA81259 and AA81304 to AA81331 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AA81322 to
CC AA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies

CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.

XX
SQ Sequence 515 BP; 110 A; 129 C; 122 G; 154 T; 0 other;

Query Match 29.0%; Score 290; DB 21; Length 515;

Best Local Similarity 100.0%; Pred. No. 1.2e-80; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0;

QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATTCACACA 771
DB 212 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATTCACACA 271

QY 772 ACATAGAGCGCGAAGCATAAGTGTAAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 831
DB 272 ACATAGAGCGCGAAGCATAAGTGTAAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 331

QY 832 CATTAATGCGTTCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTGC 891
DB 332 CATTAATGCGTTCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCGAGCTGC 391

QY 892 ATTAATGAATCGCGCAACGCGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT 951
DB 392 ATTAATGAATCGCGCAACGCGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT 451

QY 952 CTCGCTCACTAGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
DB 452 CTCGCTCACTAGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 501

RESULT 9

ABV45308
ID ABV45308 standard; cDNA; 561 BP.
XX
AC ABV45308;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 45299.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.

XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX
XX 16-MAR-2000; 2000US-189862P.
XX
XX 25-MAY-2000; 2000US-207454P.
XX
XX 09-JUN-2000; 2000US-211314P.
XX
XX 18-JUL-2000; 2000US-219007P.
XX
XX 13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
XX

PI Schlegel R, Endege WO, Monahan JB;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 8966; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 XX Sequence 561 BP; 126 A; 146 C; 159 G; 130 T; 0 other;
 SQ
 Query Match 29.0%; Score 290; DB 23; Length 561;
 Best Local Similarity 100.0%; Pred. No. 1.1e-80;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 712 CGTAATCATGTCATAGCTGTTCTCTGTGTAATTTGTAATTCGCTCACAATCCACACA 771
 124 CGTAATCATGTCATAGCTGTTCTCTGTGTAATTTGTAATTCGCTCACAATCCACACA 193
 772 ACATACGAGCGGAGCATAAAGTGAAGCTGGGGTCTTAATGAGTGAGTAACTCA 831
 184 ACATACGAGCGGAGCATAAAGTGAAGCTGGGGTCTTAATGAGTGAGTAACTCA 243
 832 CATTAATTCGCTGCGCTCACTGCGCGCTTCCAGTCGGGAACTCTGTCGCGAGCTGC 891
 244 CATTAATTCGCTGCGCTCACTGCGCGCTTCCAGTCGGGAACTCTGTCGCGAGCTGC 303
 892 ATTAATGAATCGGCAACGCGGCGGAGAGCGGTTTGGATTGGCGCTCTTCGCTT 951
 304 ATTAATGAATCGGCAACGCGGCGGAGAGCGGTTTGGATTGGCGCTCTTCGCTT 363
 952 CTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 364 CTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 413
 RESULT 10
 AAF68704
 ID AAF68704 standard; cDNA; 584 BP.
 XX
 AC AAF68704;
 XX
 DT 12-APR-2001 (first entry)
 XX
 DE Human lung tumour protein related nucleotide sequence SEQ ID NO:639.
 XX
 KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
 KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
 KW cytostatic; antisense inhibition; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200100828-A2.
 FN
 XX 04-JAN-2001.
 PD
 XX

PF 30-JUN-2000; 2000WO-US18061.
 XX
 PR 30-JUN-1999; 99US-0346492.
 PR 15-OCT-1999; 99US-0419356.
 PR 17-DEC-1999; 99US-0466867.
 PR 30-DEC-1999; 99US-0476300.
 PR 06-MAR-2000; 2000US-0519642.
 PR 22-MAR-2000; 2000US-0533077.
 PR 10-APR-2000; 2000US-0546259.
 PR 27-APR-2000; 2000US-0560406.
 PR 05-JUN-2000; 2000US-0589184.
 XX (CORI-) CORIXA CORP.
 PA
 XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
 PI Retter MW, Mannion J;
 XX WPI; 2001-071488/08.
 DR
 XX Lung tumor-associated proteins and the nucleic acids that encode them,
 PT useful for preventing, diagnosing and treating lung cancer -
 PT
 XX Claim 4; Page 360; 436pp; English.
 PS
 XX The present invention describes immunogenic portions of lung tumour-
 CC associated proteins (I) and the nucleic acids (NAs) that encode them.
 CC (I) have cytostatic activity and can be used in gene therapy, antisense
 CC inhibition and in vaccines. The NAs and the lung tumour-associated
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with their inappropriate expression,
 CC especially lung cancers. For example, the NAs may be administered to
 CC treat diseases by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of the protein by expressing inactive proteins
 CC to supplement the patient's own production of (I). Additionally, the
 CC NAs may be used to produce the lung-tumour associated protein, according
 CC to standard recombinant DNA methodology. Conversely, antisense NA
 CC molecules may be administered to down regulate protein expression by
 CC binding with the cells own genes and preventing their expression. The NA
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar NA sequences in
 CC samples, and hence which patients may be in need of treatment for lung
 CC cancer. The (I) may be used as antigens in the production of antibodies
 CC and in assays to identify modulators (agonists and antagonists) of the
 CC expression and activity of the protein. AAF68083 to AAF68878 and
 CC AAF6848 to AAF6878 represent human lung tumour protein related
 CC nucleotide and protein sequences which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 584 BP; 142 A; 159 C; 153 G; 130 T; 0 other;
 Query Match 29.0%; Score 290; DB 22; Length 584;
 Best Local Similarity 100.0%; Pred. No. 1.1e-80;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 712 CGTAATCATGTCATAGCTGTTCTCTGTGTAATTTGTAATTCGCTCACAATCCACACA 771
 37 CGTAATCATGTCATAGCTGTTCTCTGTGTAATTTGTAATTCGCTCACAATCCACACA 96
 772 ACATACGAGCGGAGCATAAAGTGAAGCTGGGGTCTTAATGAGTGAGTAACTCA 831
 97 ACATACGAGCGGAGCATAAAGTGAAGCTGGGGTCTTAATGAGTGAGTAACTCA 156
 832 CATTAATTCGCTGCGCTCACTGCGCGCTTTCAGTCGGGAACTCTGTCGCGAGCTGC 891
 157 CATTAATTCGCTGCGCTCACTGCGCGCTTTCAGTCGGGAACTCTGTCGCGAGCTGC 216
 892 ATTAATGAATCGGCAACGCGGCGGAGAGCGGTTTGGATTGGCGCTCTTCGCTT 951
 217 ATTAATGAATCGGCAACGCGGCGGAGAGCGGTTTGGATTGGCGCTCTTCGCTT 276
 952 CTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 277 CTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 326

the polynucleotide, protein or antigen-presenting cells, under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells) and inhibiting the development of a cancer in a patient (by incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with the polynucleotide, protein or antigen presenting cells that express the polynucleotide, such that T cells proliferate, administering to the patient an effective amount of the proliferated T cells, and thus inhibiting the development of a cancer in the patient. The polynucleotide, protein and cells are useful in a composition for stimulating an immune response in a patient, and for treating a cancer in a patient (particularly lung cancer). The oligonucleotide is useful for determining the presence of a cancer in a patient. The protein and oligonucleotides are useful in pharmaceutical compositions, e.g. vaccines. The polynucleotide is also useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and proteins in tumour cells. An amplified portion of the polynucleotide is useful for isolating a full-length gene from a suitable library. The present sequence is a cDNA (full length, extended or partial) isolated from a library derived from lung tumour/cancer cells. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669.

Sequence 584 BP; 142 A; 159 C; 153 G; 130 T; 0 other;

Query Match 29.0%; Score 290; DB 25; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTCTCTGTGCGAAATGTTATCCGCTCACAAATCCACACA 771
Db 37 CGTAATCATGTCATAGCTGTTCTCTGTGCGAAATGTTATCCGCTCACAAATCCACACA 96

QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGTTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGTTGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAATTCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCAGCTGC 891
Db 157 CATTAATTCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCAGCTGC 216

QY 892 ATTAATGAATCGCCCAACGCGGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCGCGTT 951
Db 217 ATTAATGAATCGCCCAACGCGGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCGCGTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 326

RESULT 13
ID ABX99895 standard; cDNA; 584 BP.

XX AC

XX ABX99895;

XX 22-MAY-2003 (first entry)

XX Lung cancer therapy and diagnosis associated cDNA #622.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-0902941.

XX

PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0468667.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 12-MAR-2000; 2000US-0533077.
PR 20-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
PR 11-JUL-2000; 2000US-0614124.
PR 19-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.

(CORI-) CORIXA CORP.

Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;

Carter D, Fanger GR, Vedwick TS, Bangur CS, McNabb A;

WPI; 2003-328427/31.

New polynucleotide, useful for preparing a composition for treating or inhibiting development of cancer, e.g. lung cancer -

Example 1; SEQ ID NO 639; 82pp; English.

The invention describes an isolated polynucleotide comprising one of 32 sequences, complement or degenerate variants of them. The polynucleotide is useful for preparing a composition e.g. a vaccine or for gene therapy, for treating or inhibiting development of cancer, e.g. lung cancer. This sequence represents a polynucleotide associated with the compositions and methods for the therapy and diagnosis of lung cancer.

Sequence 584 BP; 142 A; 159 C; 153 G; 130 T; 0 other;

Query Match 29.0%; Score 290; DB 25; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTCTCTGTGCGAAATGTTATCCGCTCACAAATCCACACA 771
Db 37 CGTAATCATGTCATAGCTGTTCTCTGTGCGAAATGTTATCCGCTCACAAATCCACACA 96

QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGTTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGTTGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAATTCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCAGCTGC 891
Db 157 CATTAATTCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCAGCTGC 216

QY 892 ATTAATGAATCGCCCAACGCGGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCGCGTT 951
Db 217 ATTAATGAATCGCCCAACGCGGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCGCGTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 326

RESULT 14
ID AAA62566 standard; cDNA; 598 BP.

XX AC

XX AAA62566;

XX 14-NOV-2000 (first entry)

XX Human breast cancer cell cDNA sequence #382.

XX

XX Human; arbitrary primer; cDNA synthesis; contig sequence construction;
KW open reading frame; ORF; low stringency; cDNA sequencing; ss.
XX Homo sapiens.
OS WO200031299-A2.
PN 02-JUN-2000.
XX 19-NOV-1999; 99WO-US27430.
XX 20-NOV-1998; 98US-0196716.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Simpson AJG, Dias Neto E, Brentani RR;
PI WPI; 2000-400100/34.
XX
XX Method for determining open reading frames of the genome of an organism
PT using primers at low stringency conditions, useful in the construction
PT of contigs or constructs of sequenced nucleic acid molecules -
XX
XX Example 7; Page 112; 113pp; English.
XX
XX The present sequence is a cDNA sequence obtained using a method for
CC determining open reading frames (ORFs) of the genome of an
CC organism. An aliquot of mRNA from human breast cancer cells was mixed
CC with a single, arbitrary primer, Moloney murine leukemia virus reverse
CC transcriptase, reverse transcriptase buffer and dNTPs. The mixture was
CC incubated under low stringency conditions to yield single stranded
CC cDNA. The same primer was then used to amplify the cDNA by PCR. Rather
CC than providing nucleotide sequence information from the non-coding
CC termini of nucleic acid molecules, the method provides information on
CC the more interesting and relevant internal portions, such as ORFs. The
CC method also permits the construction of contigs of sequenced nucleic
CC acid molecules.
XX
XX Sequence 598 BP; 136 A; 170 C; 155 G; 132 T; 5 other;
SQ
Query Match 29.0%; Score 290; DB 21; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCATAGCTGTTCTCTGTGAAATTTCTATCCGCTCACAAATCCACACA 771
DB 10 CGTAATCATGTCATAGCTGTTCTCTGTGAAATTTCTATCCGCTCACAAATCCACACA 69
QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 831
DB 70 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 129
QY 832 CATTAATGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAACCTGCTGCGAGCTGC 891
DB 130 CATTAATGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAACCTGCTGCGAGCTGC 189
QY 892 ATTAATGAATCGGCAACGCGCGGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCGCGTT 951
DB 190 ATTAATGAATCGGCAACGCGCGGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCGCGTT 249
QY 952 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
DB 250 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 299
RESULT 15
ABT12681
ID ABT12681 standard; DNA; 598 BP.
XX
AC ABT12681;
XX
DT 02-JAN-2003 (first entry)

XX Orestes sequencing method related DNA SEQ ID No 387.
DE Open reading frame; ORF; low stringency; sequencing; ORESTES; bronchitis;
XX identification; internal region; population study; heart disease;
KW Alzheimer's disease; autoimmune disease; congenital disease; human; ds.
XX Homo sapiens.
OS WO200274994-A2.
PN 26-SEP-2002.
XX 01-NOV-2001; 2001WO-US46665.
XX 07-NOV-2000; 2000US-246313P.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Simpson AJG, Dias-neto E, Brentani RR;
PI WPI; 2002-750567/81.
XX
XX Determining ORFs of the genome by contacting mRNA with a single
PT oligonucleotide primer, useful for population studies that identify
PT genes associated with diseases such as cardiovascular, autoimmune and
PT congenital diseases -
XX
XX Example 7; Page 161-162; 166pp; English.
XX
XX The invention relates to a novel method for determining open reading
CC frames (ORFs) of the genome of an organism comprising contacting
CC messenger RNA from a cell of the organism with a single, oligonucleotide
CC primer at low stringency. The method is an improved ORESTES sequencing
CC method. The methods of the present invention are useful for sequencing
CC of, essentially, the entire coding component of an organism,
CC identification of previously unknown nucleic acid molecules, assigning
CC previously identified nucleotide sequences to internal regions of genes
CC and population studies that identify genes associated with diseases such
CC as heart disease, bronchitis, Alzheimer's disease, autoimmune diseases
CC and congenital diseases. This polynucleotide represents a DNA sequence
CC that is used in the novel method of the invention.
XX
XX Sequence 598 BP; 136 A; 170 C; 155 G; 132 T; 5 other;
SQ
Query Match 29.0%; Score 290; DB 24; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCATAGCTGTTCTCTGTGAAATTTCTATCCGCTCACAAATCCACACA 771
DB 10 CGTAATCATGTCATAGCTGTTCTCTGTGAAATTTCTATCCGCTCACAAATCCACACA 69
QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 831
DB 70 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 129
QY 832 CATTAATGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAACCTGCTGCGAGCTGC 891
DB 130 CATTAATGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAACCTGCTGCGAGCTGC 189
QY 892 ATTAATGAATCGGCAACGCGCGGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCGCGTT 951
DB 190 ATTAATGAATCGGCAACGCGCGGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCGCGTT 249
QY 952 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
DB 250 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 299
Search completed; February 15, 2004, 22:40:26
Job time : 327.065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 10:22:55 ; Search time 73.6147 Seconds
(without alignments)
6001.851 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000

Perfect score: 1001

Sequence: 1 tctcagggttaaccagtggtggg.....ttcggctggggagcggtta 1001

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	30.7	5552	3	US-08-155-888-1
2	290	29.0	450	4	US-09-486-336A-2
3	290	29.0	584	4	US-09-702-705-639
4	290	29.0	584	4	US-09-738-457-639
5	290	29.0	701	2	US-08-708-702-24
6	290	29.0	701	3	US-08-708-706-24
7	290	29.0	701	4	US-09-238-471-24
8	290	29.0	713	2	US-08-708-702-26
9	290	29.0	713	3	US-08-708-706-26
10	290	29.0	713	4	US-09-238-471-26
11	290	29.0	716	3	US-08-998-416-55
12	290	29.0	717	2	US-08-708-702-25
13	290	29.0	717	3	US-08-708-706-25
14	290	29.0	717	4	US-09-238-471-25
15	290	29.0	738	1	US-08-726-462B-3
16	290	29.0	738	2	US-09-046-203-3
17	290	29.0	738	3	US-09-272-104-3
18	290	29.0	738	4	US-09-272-097-3
19	290	29.0	752	4	US-08-976-259-108
20	290	29.0	809	4	US-09-328-475C-259
21	290	29.0	848	4	US-09-328-475C-195
22	290	29.0	1024	4	US-09-328-475C-70
23	290	29.0	1738	1	US-07-971-160-9
24	290	29.0	1738	1	US-08-336-241-9
25	290	29.0	1738	2	US-08-465-273-9
26	290	29.0	1738	2	US-09-119-024-9
27	290	29.0	1738	2	US-08-417-226-9

c 28 290 29.0 1738 3 US-09-196-131-9 Sequence 9, Appli
c 29 290 29.0 1738 4 US-08-843-732-9 Sequence 9, Appli
30 290 29.0 2755 3 US-07-916-098A-7 Sequence 7, Appli
31 290 29.0 2961 3 US-08-446-935-6 Sequence 6, Appli
32 290 29.0 2973 4 US-09-402-266B-17 Sequence 11, Appli
33 290 29.0 3307 4 US-09-401-171C-11 Sequence 17, Appli
34 290 29.0 3699 3 US-08-646-538-6 Sequence 6, Appli
35 290 29.0 3699 3 US-09-503-222-6 Sequence 6, Appli
36 290 29.0 3717 2 US-08-784-208-1 Sequence 1, Appli
37 290 29.0 3789 3 US-08-075-019-8 Sequence 8, Appli
38 290 29.0 3796 4 US-09-470-681A-32 Sequence 32, Appli
39 290 29.0 3799 4 US-09-173-053-4 Sequence 4, Appli
40 290 29.0 3853 3 US-08-801-092-5 Sequence 5, Appli
41 290 29.0 3918 4 US-09-315-113-5 Sequence 5, Appli
42 290 29.0 3918 2 US-08-495-500-2 Sequence 2, Appli
43 290 29.0 3956 4 US-09-402-266B-21 Sequence 21, Appli
44 290 29.0 3974 3 US-09-026-343-33 Sequence 33, Appli
45 290 29.0 3974 3 US-09-042-105-16 Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-08-155-888-1
; Sequence 1, Application US/08155888
; Patent No. 6066823
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Sedegrah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical Res. & Dev. Cmd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,888
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
; TELEFAX: (202) 295-1022
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5552 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..755
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "promoter"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /label= CMV-IE

OTHER INFORMATION: /note= "this feature acts as a promoter for any
OTHER INFORMATION: downstream DNA sequence."
OTHER INFORMATION: /citation= ([2])
FEATURE:
NAME/KEY: CDS .2367
LOCATION: 933..2367
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /codon_start= 933
OTHER INFORMATION: /function= "protein protective against malaria"
OTHER INFORMATION: /product= "protein"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /label= IL2-OSP
OTHER INFORMATION: /citation= ([1])
PUBLICATION INFORMATION:
AUTHORS: Sedegat, Martha
AUTHORS: Hedstrom, Richard C.
AUTHORS: Hoffman, Stephen L.
TITLE: Vaccination with Plasmodium yoelii CS protein
TITLE: plasmid DNA protects against malaria
JOURNAL: Science
PUBLICATION INFORMATION:
AUTHORS: Cullen, Bryan R.
TITLE: TRANS-ACTIVATION OF HUMAN IMMUNODEFICIENCY
TITLE: VIRUS OCCURS VIA A BIMODAL MECHANISM
JOURNAL: CELL
VOLUME: 46
PAGES: 973-982
DATE: 26 SEP-1986
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 4732
US-08-155-888-1

Query Match 30.7%; Score 307; DB 3; Length 5552;
Best Local Similarity 99.3%; Pred. No. 1.9e-112; Indels 0; Gaps 0;
Matches 457; Conservative 0; Mismatches 3;
QY 171 GAGGATGGTAACAGAGTGTGTAGTTTGGAGGCCATATCTCAATTCAGACGAGTGA 230
DB 2699 GAGGATGGTAACAGAGTGTGTAGTTTGGAGGCCATATCTCAATTCAGACGAGTGA 2758
QY 231 CTCTGCTCAGAGCATGCAACCTTCCTCCTGCTGCTAGTCTAGCAGGGGATAAAGTGA 290
DB 2759 CTCTGCTCAGAGCATGCAACCTTCCTCCTGCTGCTAGTCTAGCAGGGGATAAAGTGA 2818
QY 291 AGAAGCCTGGGCTAATCAGGGGCTGCTCAGCTCCTCTAATCTGATGCTTCTATGTT 350
DB 2819 AGAAGCCTGGGCTAATCAGGGGCTGCTCAGCTCCTCTAATCTGATGCTTCTATGTT 2878
QY 351 CTCTGCTCAGAGCATGCAACCTTCCTCCTGCTGCTAGTCTAGCAGGGGATAAAGTGA 410
DB 2879 CTCTGCTCAGAGCATGCAACCTTCCTCCTGCTGCTAGTCTAGCAGGGGATAAAGTGA 2938
QY 411 CAATGAGCTGGGTGAGGCGCCGGGGAGGAGTGTGACCTTCAGACCTTGGCACTGGAGGTG 470
DB 2939 CAATGAGCTGGGTGAGGCGCCGGGGAGGAGTGTGACCTTCAGACCTTGGCACTGGAGGTG 2998
QY 471 GCCCGGAGAGGCGGCGCATCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
DB 2999 GCCCGGAGAGGCGGCGCATCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3058
QY 531 CTGAGAGCTACTGCAACTAGGCGCCACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
DB 3059 CTGAGAGCTACTGCAACTAGGCGCCACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3118
QY 591 ACCTTTGAAGAGCAGCTACAGT 630
DB 3119 ACCTTTGAAGAGCAGCTACAGT 3158

RESULT 2
US-09-486-336A-2
; Sequence 2, Application US/09486336A
; Patent No. 6566336

GENERAL INFORMATION:
APPLICANT: SUGIYAMA, Hiroshi
APPLICANT: TAO, Zhi Fu
APPLICANT: SAITO, Isao
TITLE OF INVENTION: A COMPOUND TO ALKYLATE FOR THE SPECIFIC
TITLE OF INVENTION: SEQUENCE IN DNA AND ITS METHOD OF PREPARATION
FILE REFERENCE: 49430 (71526)
CURRENT APPLICATION NUMBER: US/09/486,336A
CURRENT FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCI/JP99/01228
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 10-260710/1997
PRIOR FILING DATE: 1997-09-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 450
TYPE: DNA
ORGANISM: pUC 18
US-09-486-336A-2
Query Match 29.0%; Score 290; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.6e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 771
DB 78 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 137
QY 772 ACATACGAGCCGGAAGCATATAAGTGTAAAGCTGGGTGCTTAATGAGTGAAGTAACTCA 831
DB 138 ACATACGAGCCGGAAGCATATAAGTGTAAAGCTGGGTGCTTAATGAGTGAAGTAACTCA 197
QY 832 CATTAAATTCGTTGGCTCACTGCCCGCTTTCAGTGGGAAACCTTCGTGCCAGCTGC 891
DB 198 CATTAAATTCGTTGGCTCACTGCCCGCTTTCAGTGGGAAACCTTCGTGCCAGCTGC 257
QY 892 ATTAATGAATGCGCAACGCGCGGAGAGCGGTTTGGTATTTGGGCGCTCTTCCGCTT 951
DB 258 ATTAATGAATGCGCAACGCGCGGAGAGCGGTTTGGTATTTGGGCGCTCTTCCGCTT 317
QY 952 CTCTGCTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
DB 318 CTCTGCTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367

RESULT 3
US-09-702-705-639
; Sequence 639, Application US/09702705
; Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedwick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 639
LENGTH: 584
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-639

Query Match 29.0%; Score 290; DB 4; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.5e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGCTGTAATGTTATCCGCTCACAATCCACACA 771
DB 37 CGTAATCATGGTCATAGCTGTTTCCTGCTGTAATGTTATCCGCTCACAATCCACACA 96

QY 772 ACATACGAGCGGAGCATAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 831
DB 97 ACATACGAGCGGAGCATAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 156

QY 832 CATTAATTCGGTGTGGCTCACTGCGGCTTTCCAGTCGGGAACCTGTCTGTCAGCTGC 891
DB 157 CATTAATTCGGTGTGGCTCACTGCGGCTTTCCAGTCGGGAACCTGTCTGTCAGCTGC 216

QY 892 ATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGATTTGGCGCTCTTCCGCTT 951
DB 217 ATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGATTTGGCGCTCTTCCGCTT 276

QY 952 COTCGCTCACTGACTCGCTGCGCTCGGTCGCTTCGCGCTCGCGGAGCGGTA 1001
DB 277 COTCGCTCACTGACTCGCTGCGCTCGGTCGCTTCGCGCTCGCGGAGCGGTA 326

RESULT 4
US-09-736-457-639
; Sequence 639, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darriack
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Lijun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 639
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-639

Query Match 29.0%; Score 290; DB 4; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.5e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGCTGTAATGTTATCCGCTCACAATCCACACA 771
DB 37 CGTAATCATGGTCATAGCTGTTTCCTGCTGTAATGTTATCCGCTCACAATCCACACA 96

QY 772 ACATACGAGCGGAGCATAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 831
DB 97 ACATACGAGCGGAGCATAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 156

QY 832 CATTAATTCGGTGTGGCTCACTGCGGCTTTCCAGTCGGGAACCTGTCTGTCAGCTGC 891
DB 157 CATTAATTCGGTGTGGCTCACTGCGGCTTTCCAGTCGGGAACCTGTCTGTCAGCTGC 216

QY 892 ATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGATTTGGCGCTCTTCCGCTT 951
DB 217 ATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGATTTGGCGCTCTTCCGCTT 276

QY 952 COTCGCTCACTGACTCGCTGCGCTCGGTCGCTTCGCGCTCGCGGAGCGGTA 1001
DB 277 COTCGCTCACTGACTCGCTGCGCTCGGTCGCTTCGCGCTCGCGGAGCGGTA 326

RESULT 5
US-08-706-702-24
; Sequence 24, Application US/08706702
; Patent No. 5948614
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Deb K.
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga
; TITLE OF INVENTION: maritima and Mutants Thereof
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,702
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/689,807
; FILING DATE: 14-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,400
; FILING DATE: 02-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/576,759
; FILING DATE: 21-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,397
; FILING DATE: 02-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/525,057
; FILING DATE: 08-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
US-08-706-702-24

Query Match 29.0%; Score 290; DB 2; Length 701;
Best Local Similarity 100.0%; Pred. No. 1.5e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGCTGTAATGTTATCCGCTCACAATCCACACA 771
DB 47 CGTAATCATGGTCATAGCTGTTTCCTGCTGTAATGTTATCCGCTCACAATCCACACA 106

QY 772 ACATACGAGCGGAGCATAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 831
DB 107 ACATACGAGCGGAGCATAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 166

Query Match	29.0%;	Score	290;	DB	2;	Length	713;
Best Local Similarity	100.0%;	Pred. No.	1.4e-105;				
Matches	290;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	712	CGTAATCATGCTCATAGCTGTTTCCTGTGTGAGAAATGTTATCCGCTCACAAATCCACACA	771				
Db	81	CGTAATCATGCTCATAGCTGTTTCCTGTGTGAGAAATGTTATCCGCTCACAAATCCACACA	140				
QY	772	ACATACGAGCGGAGACATAAAGTGTAAAGCTCTGGGTGCCTTAATGAGTGAAGCTAACTCA	831				
Db	141	ACATACGAGCGGAGACATAAAGTGTAAAGCTCTGGGTGCCTTAATGAGTGAAGCTAACTCA	200				
QY	832	CATTAATTTGGGTTTGGCGTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGC	891				
Db	201	CATTAATTTGGGTTTGGCGTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGC	260				
QY	892	ATTTAATGAATCGGCCAACGCGCGGGAGAGGCGGTTTCGGTATTGGGGCGCTCTCCGCTT	951				
Db	261	ATTTAATGAATCGGCCAACGCGCGGGAGAGGCGGTTTCGGTATTGGGGCGCTCTCCGCTT	320				
QY	952	CCTCGCTCACTGAATCGCTGCGCTCGGTCGGTTCGGTCTGGCGGACGGTA	1001				
Db	321	CCTCGCTCACTGAATCGCTGCGCTCGGTCGGTTCGGTCTGGCGGACGGTA	370				

RESULT 9
US-08-706-706-26
; Sequence 26, Application US/08706706
; Patent No: 6015668
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John
; APPLICANT: Chatterjee, Deb K.
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,706
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/689,807

Sequence 55, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Sabine
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 716 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: PAG1015UP
US-08-998-416-55

Query Match 29.0%; Score 290; DB 3; Length 716;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 771
DB 206 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 265
QY 772 ACATACAGCGGAGCAATAAGTGAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 831
DB 266 ACATACAGCGGAGCAATAAGTGAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 325
QY 832 CATTAATTGCGTTCAGTCCCGCTTTCAGTCCGGGAAACCTGTCGTCGAGCTGC 891
DB 326 CATTAATTGCGTTCAGTCCCGCTTTCAGTCCGGGAAACCTGTCGTCGAGCTGC 385
QY 892 ATTAATGATTCGGCAACCGCGGGAGAGCGGTTTCGTAATGGGGCTCTTCCGCTT 951
DB 386 ATTAATGATTCGGCAACCGCGGGAGAGCGGTTTCGTAATGGGGCTCTTCCGCTT 445
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTGTTTCGCTCGCGGAGCGGTA 1001
DB 446 CCTCGCTCACTGACTCGCTCGCTCGCTGTTTCGCTCGCGGAGCGGTA 495

RESULT 12
US-08-706-702-25
Sequence 25, Application US/08706702
Patent No. 5948614
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga
TITLE OF INVENTION: maritima and Mutants Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,702
FILING DATE: 06-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2800006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: DNA (genomic)
US-08-706-702-25

Query Match 29.0%; Score 290; DB 2; Length 717;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 771
DB 85 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 144
QY 772 ACATACAGCGGAGCAATAAGTGAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 831
DB 145 ACATACAGCGGAGCAATAAGTGAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 204
QY 832 CATTAATTGCGTTCGCTCACTGCCCGCTTTCAGTCCGGGAAACCTGTCGTCGAGCTGC 891
DB 205 CATTAATTGCGTTCGCTCACTGCCCGCTTTCAGTCCGGGAAACCTGTCGTCGAGCTGC 264

QY 892 ATTAATGAATCGCCCAACGCGGAGAGCGGTTTCCGTATGGCGCTCTTCGGCTT 951
Db 265 ATTAATGAATCGCCCAACGCGGAGAGCGGTTTCCGTATGGCGCTCTTCGGCTT 324
QY 952 CCTCGCTCACTGACTGCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCT 1001
Db 325 CCTCGCTCACTGACTGCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCT 374

RESULT 13

US-08-706-706-25
; Sequence 25, Application US/08706706
; Patent No. 6015668
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,706

FILING DATE: 06-SEP-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807

FILING DATE: 14-AUG-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400

FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/370,190

FILING DATE: 01-JAN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,423

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759

FILING DATE: 21-DEC-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397

FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057

FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:
NAME: Esmord, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.2800005

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2500

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 717 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULAR TYPE: DNA (genomic)

US-08-706-706-25

Query Match 29.0%; Score 290; DB 3; Length 717;

Best Local Similarity 100.0%; Pred. NO. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGCTCATAGCTGTTTCCGTGCTGAAATTTATCCGCTCACAATTCACACA 771
Db 85 CGTAATCATGCTCATAGCTGTTTCCGTGCTGAAATTTATCCGCTCACAATTCACACA 144
QY 772 ACATACAGCGCGAAGCATATAAGCTGTAAAGCTCGGGTGCCTTAATGAGTGAAGTCA 831
Db 145 ACATACAGCGCGAAGCATATAAGCTGTAAAGCTCGGGTGCCTTAATGAGTGAAGTCA 204
QY 832 CATTAATTGCTTGGCTCACTGCCGCTTCCAGTCCGGAAACCTGTCTGCCAGCTGC 891
Db 205 CATTAATTGCTTGGCTCACTGCCGCTTCCAGTCCGGAAACCTGTCTGCCAGCTGC 264
QY 892 ATTAATGAATCGCCCAACGCGGAGAGCGGTTTCCGTATGGCGCTCTTCGGCTT 951
Db 265 ATTAATGAATCGCCCAACGCGGAGAGCGGTTTCCGTATGGCGCTCTTCGGCTT 324
QY 952 CCTCGCTCACTGACTGCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCT 1001
Db 325 CCTCGCTCACTGACTGCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCT 374

RESULT 14

US-09-238-471-25
; Sequence 25, Application US/09238471
; Patent No. 6505560
; GENERAL INFORMATION:

APPLICANT: Hughes, A. John

TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

STREET: 1100 New York Ave., N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/238,471

FILING DATE: 28-Jan-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/706,706

FILING DATE: 06-SEP-1996

APPLICATION NUMBER: US 08/689,807

FILING DATE: 14-AUG-1996

APPLICATION NUMBER: US 08/537,400

FILING DATE: 02-OCT-1995

APPLICATION NUMBER: US 08/370,190

FILING DATE: 01-JAN-1995

APPLICATION NUMBER: US 08/316,423

FILING DATE: 30-SEP-1994

APPLICATION NUMBER: US 08/576,759

FILING DATE: 21-DEC-1995

APPLICATION NUMBER: US 08/537,397

FILING DATE: 02-OCT-1995

APPLICATION NUMBER: US 08/525,057

FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Esmord, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.2800005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 717 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-238-471-25

Query Match 29.0%; Score 290; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 771
DB 85 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 144
QY 772 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 831
DB 145 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 204
QY 832 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
DB 205 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 264
QY 892 ATTATGAATCGGCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTTCCGCTT 951
DB 265 ATTATGAATCGGCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTTCCGCTT 324
QY 952 CTTCCGCTCACTGACTCGTGGCTCGGTTCGTTCCGCTGCGGCGAGCGGTA 1001
DB 325 CTTCCGCTCACTGACTCGTGGCTCGGTTCGTTCCGCTGCGGCGAGCGGTA 374

RESULT 15

US-08-726-462B-3
; Sequence 3, Application US/08726462B

; Patent No. 5800996
; GENERAL INFORMATION:
; APPLICANT: Perkin-Elmer Corporation, Applied Biosystems
; APPLICANT: Division
; TITLE OF INVENTION: ENERGY TRANSFER DYES WITH ENHANCED
; TITLE OF INVENTION: FLUORESCENCE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; ADDRESSEE: & Rosati
; STREET: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Wordperfect for Windows 6.0,
; SOFTWARE: ASCII (DOS) TEXT format
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,462B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Weitz
; REGISTRATION NUMBER: 38,362
; REFERENCE/DOCKET NUMBER: PELM4304

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 493-9300
; TELEFAX: (415) 493-6811
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-726-462B-3

Query Match 29.0%; Score 290; DB 1; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 771
DB 109 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 168
QY 772 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 831
DB 169 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 228
QY 832 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
DB 229 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 288
QY 892 ATTATGAATCGGCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTTCCGCTT 951
DB 289 ATTATGAATCGGCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTTCCGCTT 348
QY 952 CTTCCGCTCACTGACTCGTGGCTCGGTTCGTTCCGCTGCGGCGAGCGGTA 1001
DB 349 CTTCCGCTCACTGACTCGTGGCTCGGTTCGTTCCGCTGCGGCGAGCGGTA 398

Search completed: February 15, 2004, 19:10:14
Job time : 75.6147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 16:21:35 ; Search time 420.484 Seconds
(without alignments)
8769.219 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000

Perfect score: 1001

Sequence: 1 tctcaggttaacagctgggg.....tctcaggtcgagcggtta 1001

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	5283	13	US-09-921-143-36
2	291	29.1	9632	12	US-10-394-511-1
3	290	29.0	380	10	US-09-796-632-3110
4	290	29.0	380	12	US-10-057-475B-3110
5	290	29.0	380	12	US-10-154-864B-3110
6	290	29.0	380	15	US-10-040-862-3110
7	290	29.0	413	13	US-10-270-839-103
8	290	29.0	557	10	US-09-796-632-3746
9	290	29.0	557	12	US-10-057-475B-3746
10	290	29.0	557	12	US-10-154-864B-3746
11	290	29.0	557	15	US-10-040-862-3746
12	290	29.0	584	10	US-09-736-457-639
13	290	29.0	584	10	US-09-902-941-639
14	290	29.0	584	10	US-09-849-626-639
15	290	29.0	584	11	US-09-476-300-639

16	290	29.0	584	13	US-10-113-872-639	Sequence 639, App
17	290	29.0	584	15	US-10-017-754-639	Sequence 639, App
18	290	29.0	605	13	US-09-814-353-18505	Sequence 18505, A
19	290	29.0	686	10	US-09-764-868-1382	Sequence 1382, Ap
20	290	29.0	686	11	US-09-764-891-8687	Sequence 8687, Ap
21	290	29.0	686	11	US-09-764-891-9713	Sequence 9713, Ap
22	290	29.0	701	11	US-09-229-173-24	Sequence 24, Appl
23	290	29.0	701	11	US-10-285-696-24	Sequence 24, Appl
24	290	29.0	713	11	US-09-229-173-26	Sequence 26, Appl
25	290	29.0	713	13	US-10-285-696-26	Sequence 26, Appl
26	290	29.0	717	11	US-09-229-173-25	Sequence 25, Appl
27	290	29.0	717	13	US-10-285-696-25	Sequence 25, Appl
28	290	29.0	752	9	US-09-956-004-108	Sequence 108, App
29	290	29.0	803	13	US-10-257-826A-17	Sequence 17, Appl
30	290	29.0	808	15	US-10-198-846-6873	Sequence 6873, Ap
31	290	29.0	809	14	US-10-202-193-259	Sequence 259, App
32	290	29.0	848	14	US-10-202-193-195	Sequence 195, App
33	290	29.0	1024	14	US-10-202-193-70	Sequence 70, Appl
34	290	29.0	1091	12	US-09-764-875-184	Sequence 184, App
35	290	29.0	1092	10	US-09-764-868-232	Sequence 232, App
36	290	29.0	1092	11	US-09-764-891-1853	Sequence 1853, Ap
37	290	29.0	1092	12	US-09-764-875-482	Sequence 482, App
38	290	29.0	1355	9	US-09-764-870-182	Sequence 182, App
39	290	29.0	1355	9	US-09-764-853-404	Sequence 404, App
40	290	29.0	1355	12	US-10-158-057-165	Sequence 165, App
41	290	29.0	1355	15	US-10-125-540-182	Sequence 182, App
42	290	29.0	1355	15	US-10-103-313-273	Sequence 273, App
43	290	29.0	1738	11	US-09-836-169-9	Sequence 9, Appl
44	290	29.0	1738	13	US-10-152-994-9	Sequence 9, Appl
45	290	29.0	2074	15	US-10-106-698-350	Sequence 350, App

ALIGNMENTS

RESULT 1
US-09-921-143-36
; Sequence 36, Application US/09921143
; Publication No: US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P6
; CURRENT APPLICATION NUMBER: US/09/921.143
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 5283
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-36

Query Match	100.0%	Score 1001;	DB 13;	Length 5283;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1001;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TCTCAGCTAACAGTGGGGGCGACATTTCTGCGGAGCTAGACATATGAACATGCT	60	
Db	2000	TCTCAGCTAACAGTGGGGGCGACATTTCTGCGGAGCTAGACATATGAACATGCT	2059	
Qy	61	AGTGCAGAGAGAGTGAAGATCTTCTTAAGTCTCTAGTGTGGTGGGCTAG	120	
Db	2060	AGTGCAGAGAGAGTGAAGATCTTCTTAAGTCTCTAGTGTGGTGGGCTAG	2119	
Qy	121	GCCCCAGGATAGTACCTATTGGGGAGCCCATAGACACTGCACTGACGAGGATGCT	180	
Db	2120	GCCCCAGGATAGTACCTATTGGGGAGCCCATAGACACTGCACTGACGAGGATGCT	2179	
Qy	181	AACAGATGTAGTGGTTTGGAGCCCATATGTCCTATCATGACCTGACTGCTTCTCAC	240	

SEQ ID NO 3110
LENGTH: 380
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-3110

Query Match 29.0%; Score 290; DB 10; Length 380;
Best Local Similarity 100.0%; Pred No. 1.2e-143; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CTTAATCATGTGTATAGCTGTTTCCCTGTGTGAATTTATCCCTCACAATTCACACA 771
Db 16 CTTAATCATGTGTATAGCTGTTTCCCTGTGTGAATTTATCCCTCACAATTCACACA 75

QY 772 ACATACGAGCGGGAAGCATAAAGCTGTAAAGCTGGGGTGCCTAATGAGTGAGTAACTCA 831
Db 76 ACATACGAGCGGGAAGCATAAAGCTGTAAAGCTGGGGTGCCTAATGAGTGAGTAACTCA 135

QY 832 CATTAAATTTGGCTTCCGCTCACTCCCGCTTTCAGTCCGGAAACCTGTGCTGCCAGCTGC 891
Db 136 CATTAAATTTGGCTTCCGCTCACTCCCGCTTTCAGTCCGGAAACCTGTGCTGCCAGCTGC 195

QY 892 ATTAATGATCGGCAACGCGGCGGAGAGCGGCTTTCGCTATTGGGCGCTTCCGCTT 951
Db 196 ATTAATGATCGGCAACGCGGCGGAGAGCGGCTTTCGCTATTGGGCGCTTCCGCTT 255

QY 952 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
Db 256 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 305

RESULT 4

US-10-057-475B-3110
Sequence 3110, Application US/10057475B
Publication No. US20040002069A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordenez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Hematological Malignancies
FILE REFERENCE: 014058-014402US
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3110
LENGTH: 380
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-475B-3110

Query Match 29.0%; Score 290; DB 12; Length 380;
Best Local Similarity 100.0%; Pred No. 1.2e-143; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CTTAATCATGTGTATAGCTGTTTCCCTGTGTGAATTTATCCCTCACAATTCACACA 771
Db 16 CTTAATCATGTGTATAGCTGTTTCCCTGTGTGAATTTATCCCTCACAATTCACACA 75

QY 772 ACATACGAGCGGGAAGCATAAAGCTGTAAAGCTGGGGTGCCTAATGAGTGAGTAACTCA 831
Db 76 ACATACGAGCGGGAAGCATAAAGCTGTAAAGCTGGGGTGCCTAATGAGTGAGTAACTCA 135

QY 832 CATTAAATTTGGCTTCCGCTCACTCCCGCTTTCAGTCCGGAAACCTGTGCTGCCAGCTGC 891
Db 136 CATTAAATTTGGCTTCCGCTCACTCCCGCTTTCAGTCCGGAAACCTGTGCTGCCAGCTGC 195

QY 892 ATTAATGATCGGCAACGCGGCGGAGAGCGGCTTTCGCTATTGGGCGCTTCCGCTT 951
Db 196 ATTAATGATCGGCAACGCGGCGGAGAGCGGCTTTCGCTATTGGGCGCTTCCGCTT 255

QY 952 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
Db 256 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 305

RESULT 5

US-10-154-884B-3110
Sequence 3110, Application US/10154884B
Publication No. US20040005561A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc W.
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013521US
CURRENT FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 11290
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3110
LENGTH: 380
TYPE: DNA
ORGANISM: Homo sapiens

US-10-154-884B-3110

Query Match 29.0%; Score 290; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACA 771
Db 16 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACA 75

QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 831
Db 76 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 135

QY 832 CATTAATTTGGCTTTCGCTCACTGCCCGCTTTCAGTTCGCGAAACCTGTCGCGAGCTGC 891
Db 136 CATTAATTTGGCTTTCGCTCACTGCCCGCTTTCAGTTCGCGAAACCTGTCGCGAGCTGC 195

QY 892 ATTAATGAATCGGCAACGCGGGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCTT 951
Db 196 ATTAATGAATCGGCAACGCGGGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCTT 255

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 256 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 305

RESULT 6

US-10-040-862-3110
; Sequence 310, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-01352005
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3110
; LENGTH: 380
; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-040-862-3110

Query Match 29.0%; Score 290; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACA 771
Db 16 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACA 75

QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 831
Db 76 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 135

QY 832 CATTAATTTGGCTTTCGCTCACTGCCCGCTTTCAGTTCGCGAAACCTGTCGCGAGCTGC 891
Db 136 CATTAATTTGGCTTTCGCTCACTGCCCGCTTTCAGTTCGCGAAACCTGTCGCGAGCTGC 195

QY 892 ATTAATGAATCGGCAACGCGGGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCTT 951
Db 196 ATTAATGAATCGGCAACGCGGGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCTT 255

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 256 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 305

RESULT 7

US-10-270-839-103
; Sequence 103, Application US/10270839
; Publication No. US20030143586A1
; GENERAL INFORMATION:
; APPLICANT: Chao, Qimin
; APPLICANT: Grasso, Luigi
; APPLICANT: Sass, Philip M.
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Genetic Hypermutability of Plants for Gene Discovery and Diagn
; FILE REFERENCE: AG000205 (MOR-0133)
; CURRENT APPLICATION NUMBER: US/10/270,839
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/328,750
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Conyza sp.
US-10-270-839-103

Query Match 29.0%; Score 290; DB 13; Length 413;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACA 771
Db 83 CGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACA 142

QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 831
Db 143 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 202

QY 832 CATTAATTTGGCTTTCGCTCACTGCCCGCTTTCAGTTCGCGAAACCTGTCGCGAGCTGC 891
Db 203 CATTAATTTGGCTTTCGCTCACTGCCCGCTTTCAGTTCGCGAAACCTGTCGCGAGCTGC 262

QY 892 ATTAATGAATCGGCAACGCGGGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCTT 951
Db 263 ATTAATGAATCGGCAACGCGGGGAGAGCGGTTTGGTATTTGGGCGCTTTCGCTT 322

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
|||

QY 892 ATTAATGAATCGCCAACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCT:

Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	712	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCAAAATTCACACA	771
Db	37	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCAAAATTCACACA	96
QY	772	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTAACTCA	831
Db	97	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTAACTCA	156
QY	832	CATTAAATTCGGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGCCAGCTGC	891
Db	157	CATTAAATTCGGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGCCAGCTGC	216
QY	892	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	951
Db	217	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	276
QY	952	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	1001
Db	277	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	326

RESULT 15

US-09-476-300-639
; Sequence 639, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 639
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-639

Query Match 29.0%; Score 290; DB 11; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	712	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCAAAATTCACACA	771
Db	37	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGCTCAAAATTCACACA	96
QY	772	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTAACTCA	831
Db	97	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTAACTCA	156
QY	832	CATTAAATTCGGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGCCAGCTGC	891
Db	157	CATTAAATTCGGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGCCAGCTGC	216
QY	892	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	951
Db	217	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	276
QY	952	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	1001
Db	277	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	326

Search completed: February 15, 2004, 22:26:48
Job time : 423.484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 21:50:21 ; Search time 2636.53 Seconds
(without alignments)
9227.587 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000
Perfect score: 1001
Sequence: 1 ttcaggctaaccagtgagg.....ttcggtcggcgagcgga 1001

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_esttc.*
9: gb_est1.*
10: gb_est2.*
11: gb_est3.*
12: gb_est4.*
13: gb_est5.*
14: gb_est6.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	29.0	362	12	BI940910
2	290	29.0	391	12	BI675829
3	290	29.0	391	12	BI937546
4	290	29.0	398	9	AU224539

5	290	29.0	401	12	BI675832
6	290	29.0	418	10	BI675832
7	290	29.0	418	12	BI941233
8	290	29.0	420	28	AZ254218
9	290	29.0	424	12	BI937686
10	290	29.0	424	12	BI938078
11	290	29.0	424	12	BI938097
12	290	29.0	424	12	BI938182
13	290	29.0	424	12	BI938185
14	290	29.0	424	12	BI938325
15	290	29.0	424	12	BI938341
16	290	29.0	424	12	BI938341
17	290	29.0	427	12	BI675820
18	290	29.0	427	12	BI937679
19	290	29.0	427	12	BI938087
20	290	29.0	428	12	BI937859
21	290	29.0	428	12	BI675823
22	290	29.0	431	12	BI937767
23	290	29.0	436	28	AQ015737
24	290	29.0	438	12	BI937553
25	290	29.0	439	12	BI938342
26	290	29.0	440	12	BI941001
27	290	29.0	442	12	BI941064
28	290	29.0	446	12	BI937775
29	290	29.0	448	12	BI938167
30	290	29.0	448	12	BI938314
31	290	29.0	449	12	BI938324
32	290	29.0	450	12	BI937961
33	290	29.0	450	12	BI938090
34	290	29.0	451	14	CB865735
35	290	29.0	456	9	AU224531
36	290	29.0	456	12	BI937790
37	290	29.0	458	12	BI84963
38	290	29.0	459	14	CB876419
39	290	29.0	460	9	AW790424
40	290	29.0	461	28	B83369
41	290	29.0	462	12	BI937801
42	290	29.0	462	12	BI937979
43	290	29.0	463	14	CB859029
44	290	29.0	467	28	B82861
45	290	29.0	468	29	TA176D09P

ALIGNMENTS

RESULT 1

BI940910

LOCUS

DEFINITION

CATALASE ; mRNA sequence.

ACCESSION

BI940910

VERSION

BI940910.1

KEYWORDS

EST.

SOURCE

Xenopus laevis

ORGANISM

Xenopus laevis

(African clawed frog)

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.

1 (bases 1 to 362)

Clifton, S., Johnson, S., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Willson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

BI940910 362 bp mRNA linear EST 18-OCT-2001
df08b08.x1 Wellcome CRC PRN3 St13 17 egg animal cap Xenopus laevis
cDNA clone IMAGE:3556574 3' similar to SW:CAT_A_CAWJE_Q59296
CATALASE ; mRNA sequence.

BI940910 GI:16255382

EST.

Xenopus laevis

Xenopus laevis

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.

1 (bases 1 to 362)

Clifton, S., Johnson, S., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Willson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

BI675832 dag61901.
BI361564 pot-1-64
BI941233 df56q12.x
AZ254218 G0029P.R
BI937686 de13f03.x
BI938078 de33b12.x
BI938097 de33f05.x
BI938182 de35g11.x
BI938185 de35h10.x
BI938325 de41b11.x
BI938341 de41f04.x
BI940816 de22a02.x
BI675820 dag61a04.
BI937679 de13e01.x
BI938087 de33c10.x
BI937859 de18a02.x
BI675823 dag61c01.
BI937767 de35g11.x
AQ015737 Ctr-HSP-2
BI937553 dd92e12.Y
BI938342 de41f05.x
BI941001 df11b08.x
BI941064 df13b05.x
BI937775 de16a01.x
BI938167 de35e11.x
BI938314 de41a03.x
BI938324 de41b09.x
BI937961 de29b09.x
BI938090 de33d06.x
CB865735 HD10107W
AU224531 AU224531
BI937790 de16c03.x
BI84963 fl17c01.Y
CB876419 HX11D14W
AW790424 C01743-F
B83369 RPC111-13C2
BI937801 de16d10.x
BI937979 de29e08.x
CB859029 HI09B01W
B82861 RPC111-18D2
AL475038 T. brucei

Email: est@watson.wustl.edu
 Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
 (Wellcome/CRC Institute). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 350.

FEATURES

source

1. .362
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3556574"
 /tissue_type="egg, subtracted by stage 13-17 animal cap"
 /lab_host="PH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 St13 17 egg animal cap"
 /notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is subtracted
 and was constructed by N. Garrett, E. Bellefroid, and A.M.
 Zorn, (Wellcome/CRC Institute)."
 77 a 93 c 97 g 95 t

BASE COUNT
 ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.3e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATGTTATCCGCTCAACATTCACACA 771
 Db 69 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATGTTATCCGCTCAACATTCACACA 128
 QY 772 ACATACGAGCGGAGCATTAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAAGTCACTCA 831
 Db 129 ACATACGAGCGGAGCATTAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAAGTCACTCA 188
 QY 832 CATTAAATGCTGCTCACTGCCGCTCTTCCAGTCGGGAACTGTCGTCGAGCTGC 891
 Db 189 CATTAAATGCTGCTCACTGCCGCTCTTCCAGTCGGGAACTGTCGTCGAGCTGC 248
 QY 892 ATTAATGAATCGGCCAACCGCGGAGAGCGGTTTTCGTTATGGCGCTTCCGCTT 951
 Db 249 ATTAATGAATCGGCCAACCGCGGAGAGCGGTTTTCGTTATGGCGCTTCCGCTT 308
 QY 952 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
 Db 309 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 358

RESULT 2

Bi937546

LOCUS

DEFINITION

Bi937546 391 bp mRNA linear EST 17-SEP-2001
 dag61e05.y1 Wellcome CRC PRN3 St19 26 egg animal cap Xenopus laevis
 cDNA clone IMAGE:4785896 5' similar to TR:Q44068 Q44068
 ALPHA-HEMOLYSIN. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis

(African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 391)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.

, Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person

, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,

Waterston, R., and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

REFERENCE

AUTHORS

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.

, Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person

, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,

Waterston, R., and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
 (Wellcome/CRC Institute). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40RP from Gibco
 High quality sequence stop: 273.

FEATURES

source

1. .391
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:4785896"
 /tissue_type="egg, subtracted by stage 19-26 animal cap"
 /lab_host="PH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 St19 26 egg animal cap"
 /notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is subtracted
 and was constructed by N. Garrett and A.M. Zorn,
 (Wellcome/CRC Institute)."
 87 a 98 c 106 g 100 t

BASE COUNT
 ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 391;
 Best Local Similarity 100.0%; Pred. No. 4.4e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATGTTATCCGCTCAACATTCACACA 771
 Db 36 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATGTTATCCGCTCAACATTCACACA 95
 QY 772 ACATACGAGCGGAGCATTAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAAGTCACTCA 831
 Db 96 ACATACGAGCGGAGCATTAAGTGTAAAGCTCGGGTGCCTTAATGAGTGAAGTCACTCA 155
 QY 832 CATTAAATGCTGCTCACTGCCGCTCTTCCAGTCGGGAACTGTCGTCGAGCTGC 891
 Db 156 CATTAAATGCTGCTCACTGCCGCTCTTCCAGTCGGGAACTGTCGTCGAGCTGC 215
 QY 892 ATTAATGAATCGGCCAACCGCGGAGAGCGGTTTTCGTTATGGCGCTTCCGCTT 951
 Db 216 ATTAATGAATCGGCCAACCGCGGAGAGCGGTTTTCGTTATGGCGCTTCCGCTT 275
 QY 952 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
 Db 276 CTTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 325

RESULT 3

Bi937546

LOCUS

DEFINITION

Bi937546 391 bp mRNA linear EST 18-OCT-2001
 dd92b12.y1 Wellcome CRC pcdNAI egg Xenopus laevis cDNA clone
 IMAGE:3431758 5' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis

(African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 391)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.

, Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person

, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,

Waterston, R., and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

TITLE Waterston, R. and Wilson, R.
JOURNAL WashU Xenopus EST project, 1999
COMMENT Unpublished
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon, (Wellcome/CRC Institute). DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LUNL at: info@image.lunl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 367.

FEATURES

source
1. 391
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:1341758"
/tissue_type="egg"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC pcDNA1 egg"
/note="vector: pcDNA1; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Library was constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute)."
93 a 102 c 105 g 89 t 2 others

BASE COUNT

ORIGIN
Query Match 29.0%; Score 290; DB 12; Length 391;
Best Local Similarity 100.0%; Pred. No. 4.4e-136;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCACAAATCCACACA 771
Db 12 CGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCACAAATCCACACA 71
QY 772 ACATACGAGCGGAGCATTAAGTAAAGCTGGGTGCTTAATGATGAGTAACTCA 831
Db 72 ACATACGAGCGGAGCATTAAGTAAAGCTGGGTGCTTAATGATGAGTAACTCA 131
QY 832 CATTAATTCGCTCACTGCGCTTTCAGTCGGAAACCTGTGTCGACGCTGC 891
Db 132 CATTAATTCGCTCACTGCGCTTTCAGTCGGAAACCTGTGTCGACGCTGC 191
QY 892 ATTAATGAATCGGCAACGCGCGGAGAGCGGTTTGCGTATTTGGCGCTCTTCGCTT 951
Db 192 ATTAATGAATCGGCAACGCGCGGAGAGCGGTTTGCGTATTTGGCGCTCTTCGCTT 251
QY 952 CCTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 252 CCTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

RESULT 4

AU224539
LOCUS AU224539 Ipomoea trifida stigma 398 bp mRNA linear EST 21-OCT-2002
DEFINITION AU224539 Ipomoea trifida stigma Ipomoea trifida cDNA clone ISM-0348 mRNA sequence.

ACCESSION

VERSION AU224539.1 GI:24207512

KEYWORDS

SOURCE EST.
ORGANISM Ipomoea trifida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Convolvulaceae; Ipomoea.

REFERENCE

1 (bases 1 to 398)
Tsuchiya, T., Ando, A., Ogawa, C., Futagami, K., Watase, T. and Koyama

TITLE

JOURNAL

COMMENT

, Y.

Expressed Sequence Tags from Reproductive Organs of Ipomoea trifida
cDNA Libraries
Unpublished
Contact: Tohru Tsuchiya
Faculty of Bioresources
Mie University
1515 Kamihama, Tsu, Mie 514-8507, Japan
Tel: 81-59-231-9515
Fax: 81-59-231-9515
Email: tsuchiya@bio.mie-u.ac.jp.

FEATURES

source
1. 398
/organism="Ipomoea trifida"
/mol_type="mRNA"
/db_xref="taxon:35884"
/clone="ISM-0348"
/tissue_type="stigma"
/dev_stage="tri-nucleate pollen stage"
/clone_lib="Ipomoea trifida stigma"
80 a 107 c 109 g 102 t

BASE COUNT

ORIGIN
Query Match 29.0%; Score 290; DB 9; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.4e-136;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCACAAATCCACACA 771
Db 67 CGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCACAAATCCACACA 126
QY 772 ACATACGAGCGGAGCATTAAGTAAAGCTGGGTGCTTAATGATGAGTAACTCA 831
Db 127 ACATACGAGCGGAGCATTAAGTAAAGCTGGGTGCTTAATGATGAGTAACTCA 186
QY 832 CATTAATTCGCTCACTGCGCTTTCAGTCGGAAACCTGTGTCGACGCTGC 891
Db 187 CATTAATTCGCTCACTGCGCTTTCAGTCGGAAACCTGTGTCGACGCTGC 246
QY 892 ATTAATGAATCGGCAACGCGCGGAGAGCGGTTTGCGTATTTGGCGCTCTTCGCTT 951
Db 247 ATTAATGAATCGGCAACGCGCGGAGAGCGGTTTGCGTATTTGGCGCTCTTCGCTT 306
QY 952 CCTCGCTCACTGACTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCT 1001
Db 307 CCTCGCTCACTGACTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTT 356

RESULT 5

BI675832
LOCUS BI675832 401 bp mRNA linear EST 17-SEP-2001
DEFINITION dag61g01.y1 Wellcome CRC PRN3 St19 26 egg animal cap Xenopus laevis cDNA clone IMAGE:4785984 5' similar to TR:Q44068 ALPHA-HEMOLYSIN. ; mRNA sequence.

ACCESSION

VERSION BI675832.1 GI:15628739

KEYWORDS

SOURCE EST.
ORGANISM Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

REFERENCE

1 (bases 1 to 401)
Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

WashU Xenopus EST project, 1999
Unpublished
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
(Wellcome/CRC Institute). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 291.

FEATURES

source

Location/Qualifiers
1. .401
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4785984"
/tissue_type="egg, subtracted by stage 19-26 animal cap"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC PRN3 St19 26 egg animal cap"
/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Mieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
88 a 102 c 109 g 102 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 401;
Best Local Similarity 100.0%; Pred. No. 4.4e-136; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 771
Db 46 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 105
QY 772 ACATAGAGCGGGAAGCATAAAGTGAAGCTGGGTCCTTAAGTAGTGACTTAACCTCA 831
Db 106 ACATAGAGCGGGAAGCATAAAGTGAAGCTGGGTCCTTAAGTAGTGACTTAACCTCA 165
QY 832 CATTAATTCGGTTCGCTCACTGCCGCTTCCAGTCGGGAACCTGTCGCCAGCTGC 891
Db 166 CATTAATTCGGTTCGCTCACTGCCGCTTCCAGTCGGGAACCTGTCGCCAGCTGC 225
QY 892 ATTAATGAATCGCCAAACGGCGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 951
Db 226 ATTAATGAATCGCCAAACGGCGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 285
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 286 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 335

RESULT 6

BF361564

LOCUS

DEFINITION
BF361564 418 bp mRNA linear EST 24-NOV-2000
pot-1-64 Differentially expressed cDNA library of optic tectum of
pigeon with removal of left retina Columba livia cDNA, mRNA
sequence.

ACCESSION

BF361564

VERSION

BF361564.1

KEYWORDS

EST.

SOURCE

Columba livia (domestic pigeon)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.

1 (bases 1 to 418)

REFERENCE

AUTHORS

TITLE

A profile of differentially expressed genes in p10 pigeon optic
tecta with removal of left retina after posthatching first day by
SSH

JOURNAL

COMMENT

Unpublished
Contact: Lixia Lu
Department of Biochemistry

Shanghai Tiedao University Medical School
Gonghexin Rd.1238, Shanghai, 200070, China
Email: jinxia@public.sta.net.cn.

FEATURES

source

Location/Qualifiers
1. .418
/organism="Columba livia"
/mol_type="mRNA"
/db_xref="taxon:8932"
/sex="female and male"
/tissue_type="optic tecta"
/dev stage="10 day postnatal"
/clone_lib="Differentially expressed cDNA library of optic
tectum of pigeon with removal of left retina"
/note="Differentially expressed cDNA library of optic
tectum of pigeon with removal of left retina"
79 a 120 c 118 g 101 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 290; DB 10; Length 418;
Best Local Similarity 100.0%; Pred. No. 4.5e-136; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 771
Db 96 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 155
QY 772 ACATAGAGCGGGAAGCATAAAGTGAAGCTGGGTCCTTAAGTAGTGACTTAACCTCA 831
Db 156 ACATAGAGCGGGAAGCATAAAGTGAAGCTGGGTCCTTAAGTAGTGACTTAACCTCA 215
QY 832 CATTAATTCGGTTCGCTCACTGCCGCTTCCAGTCGGGAACCTGTCGCCAGCTGC 891
Db 216 CATTAATTCGGTTCGCTCACTGCCGCTTCCAGTCGGGAACCTGTCGCCAGCTGC 275
QY 892 ATTAATGAATCGCCAAACGGCGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 951
Db 276 ATTAATGAATCGCCAAACGGCGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 335
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 336 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 385

RESULT 7

BF941233

LOCUS

DEFINITION
BF941233 418 bp mRNA linear EST 18-OCT-2001
df26g12.x1 Wellcome CRC PRN3 head Xenopus laevis cDNA clone
IMAGE:3579838 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;,
mRNA sequence.

ACCESSION

BF941233

VERSION

BF941233.1

KEYWORDS

EST.

SOURCE

Xenopus laevis (African clawed frog)

ORGANISM

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 418)

REFERENCE

AUTHORS

TITLE

WashU Xenopus EST project, 1999

JOURNAL

COMMENT

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by N. Garrett, K. Ryan and A.M. Zorn,
(Wellcome/CRC Institute). DNA Sequencing by: Washington University

Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LINL at: info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40up from Gibco
 High quality sequence stop: 336.
 Location/Qualifiers

FEATURES

source
 1..418
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3579838"
 /lab_host="DH10B (phage-resistant)"
 /tissue_type="head, stage 30"
 /clone_lib="Wellcome CRC PRN3 head"
 /notes="vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library was constructed by N. Garrett, K. Ryan and A.M. Zorn, (Wellcome/CRC Institute)."
 BASE COUNT 96 a 105 c 113 g 104 t
 ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 418;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGTCATAGCTCTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 771
 Db 69 CGTAATCATGTCATAGCTCTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 128
 QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGTAACTCA 831
 Db 129 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGTAACTCA 188
 QY 832 CATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
 Db 189 CATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
 QY 892 ATTAATGAATCGGCGCAACGCGGAGAGCGGGTTCGTTATGCGGCTCTTCGCGCTT 951
 Db 249 ATTAATGAATCGGCGCAACGCGGAGAGCGGGTTCGTTATGCGGCTCTTCGCGCTT 308
 QY 952 CTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001
 Db 309 CTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

RESULT 8
 AZ254218 420 bp DNA linear GSS 21-JUN-2000
 LOCUS
 DEFINITION
 cG0029F RFLP sequences of cowpea, Vigna unguiculata Vigna
 unguiculata genomic, genomic survey sequence.
 ACCESSION
 AZ254218
 VERSION
 AZ254218.1 GI:8602455
 KEYWORDS
 GSS.
 SOURCE
 Vigna unguiculata (cowpea)
 Vigna unguiculata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 1 (bases 1 to 420)
 Denny,R., Danesh,D., Mudge,J., Cooper,A., Larson,K., Fatokun,C. and
 Young,N.D.
 RFLP sequences of cowpea, Vigna unguiculata
 Unpublished
 Contact: Young Nevin D
 Department of Plant Pathology
 University of Minnesota
 435 Borlaug Hall, 1991 Upper Buford Circle, St. Paul MN 55108, USA
 Tel: 612 625 2225
 Fax: 612 625 9728
 Email: nevin@tc.umn.edu

TITLE
 JOURNAL
 COMMENT

Sequence of mapped RFLP marker cG0029 on linkage group vi of
 cowpea. For more information, see BeanGenes at:
 http://ars-genome.cornell.edu/cgi-bin/WebAce/webace?db=beanGenes&cl
 ass=Locus. Please see as authorities for mapping/naming: Fatokun,
 C.A., Danesh, D., Young, N.D. (1993) RFLP linkage map for cowpea,
 (Vigna unguiculata (L.) Walp.) in: Genetic Maps, 1992. S. J. O'Brien
 , ed. Cold Spring Harbor Press, Cold Spring Harbor, NY, pp.
 6.256-6.258 - and - Menancio-Hautea, D., Kumar, L., Danesh, D.,
 Young, N.D. (1993) RFLP linkage map for mungbean (Vigna radiata (L.
) Wilczek) in: Genetic Maps, 1992. S. J. O'Brien, ed. Cold Spring
 Harbor Press, Cold Spring Harbor, NY, pp. 6.259-6.260.
 Insert Length: 400 Std Error: 0.00
 Seq primer: M13F
 Class: RFLP probe.

FEATURES

Location/Qualifiers
 1..420
 /organism="Vigna unguiculata"
 /mol_type="genomic DNA"
 /cultivar="California Blackeye #5"
 /db_xref="taxon:3917"
 /tissue_type="Hypocotyl and roots"
 /dev_stage="Sprouts"
 /clone_lib="RFLP sequences of cowpea, Vigna unguiculata"
 /note="vector: pUC 18; Site 1: Pst I; DNA was digested
 with Pst I, size separated by sucrose gradient
 centrifugation and the fraction between 500-3000 base
 pairs ligated into the vector using standard method."
 BASE COUNT 98 a 104 c 112 g 90 t 16 others
 ORIGIN

Query Match 29.0%; Score 290; DB 28; Length 420;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 771
 Db 20 CGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 79
 QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGTAACTCA 831
 Db 80 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGTAACTCA 139
 QY 832 CATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
 Db 140 CATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199
 QY 892 ATTAATGAATCGGCGCAACGCGGAGAGCGGGTTCGTTATGCGGCTCTTCGCGCTT 951
 Db 200 ATTAATGAATCGGCGCAACGCGGAGAGCGGGTTCGTTATGCGGCTCTTCGCGCTT 259
 QY 952 CTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001
 Db 260 CTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309

RESULT 9
 BI937686 424 bp mRNA linear EST 18-OCT-2001
 LOCUS
 DEFINITION
 del3f03.x1 Wellcome CRC PRN3 dorsa lip Xenopus laevis cDNA clone
 IMAGE:3438076 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;
 mRNA sequence.
 ACCESSION
 BI937686
 VERSION
 BI937686.1 GI:16252159
 KEYWORDS
 EST.
 SOURCE
 Xenopus laevis (African clawed frog)
 ORGANISM
 Xenopus laevis
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 424)
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
 , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.
 1. (bases 1 to 424)
 Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
 WashU Xenopus EST project, 1999
 Unpublished
 Contact: Sandy Clifton, Ph.D.
 WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
 Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: infoimage.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 379.
 Location/Qualifiers
 1. .424
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3473792"
 /issue_type="dorsal lip"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 dorsal lip"
 /notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library was constructed
 by A.M. Zorn (Wellcome/CRC Institute)."

FEATURES

source

BASE COUNT 98 a 106 c 114 g 106 t
 ORIGIN
 Query Match 29.0%; Score 290; DB 12; Length 424;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGGTCATAGCTTTCCCTGTTGAATTTGTTATCCGCTCACAATTCACACA 771
 Db 69 CGTAATCATGGTCATAGCTTTCCCTGTTGAATTTGTTATCCGCTCACAATTCACACA 128
 QY 772 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 831
 Db 129 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 188
 QY 832 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCGCAGCTGC 891
 Db 189 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCGCAGCTGC 248
 QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCTT 951
 Db 249 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCTT 308
 QY 952 CCTCGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGGCGGAGCGGTA 1001
 Db 309 CCTCGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGGCGGAGCGGTA 358

RESULT 12
 B1938182
 LOCUS
 DEFINITION
 de35g11.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone
 IMAGE:3473792 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;
 mRNA sequence.
 ACCESSION
 B1938182
 VERSION
 B1938182.1 GI:16252654
 KEYWORDS
 EST.

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)
 Xenopus laevis

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.

REFERENCE

AUTHORS

1. (bases 1 to 424)
 Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
 WashU Xenopus EST project, 1999
 Unpublished

TITLE

JOURNAL

COMMENT

Contact: Sandy Clifton, Ph.D.
 WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
 Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: infoimage.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 309.
 Location/Qualifiers

FEATURES

source

1. .424
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3473876"
 /issue_type="dorsal lip"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 dorsal lip"
 /notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library was constructed
 by A.M. Zorn (Wellcome/CRC Institute)."

BASE COUNT 99 a 106 c 113 g 106 t
 ORIGIN
 Query Match 29.0%; Score 290; DB 12; Length 424;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTTTCCCTGTTGAATTTGTTATCCGCTCACAATTCACACA 771
 Db 69 CGTAATCATGGTCATAGCTTTCCCTGTTGAATTTGTTATCCGCTCACAATTCACACA 128
 QY 772 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 831
 Db 129 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 188
 QY 832 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCGCAGCTGC 891
 Db 189 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCGCAGCTGC 248
 QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCTT 951
 Db 249 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCTT 308
 QY 952 CCTCGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGGCGGAGCGGTA 1001
 Db 309 CCTCGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGGCGGAGCGGTA 358

RESULT 13

B1938185

LOCUS

DEFINITION

B1938185
 de35h10.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone
 IMAGE:3473922 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;
 mRNA sequence.

RESULT 15	BI938341	424 bp	linear	EST 18-OCT-2001	
LOCUS	BI938341				
DEFINITION	deaf1f04.x1 Wellcome CRC pRN3 dorsal lip <i>Xenopus laevis</i> cDNA clone				
	IMAGE:3474558 3' similar to TR:Q44068 ALPHA-HEMOLYSIN. i,				
	mRNA sequence.				
Db	309	CCTCGCTCACTGACTCGCTCGGTCGGTTCGGCTGCGGAGCGGTA	358		
		Search completed: February 16, 2004, 03:15:45			
		Job time : 2640.53 secs			

	Query Match	29.0%; Score 290; DB 12; Length 424;
	Best Local Similarity	100.0%; Pred. No. 4.5e-136;
	Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	712	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAAAATCCACACA 771
Db	69	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAAAATCCACACA 128
QY	772	ACATACGAGCCGGAGCATAAAGTGTAAAGCCTCGGGTGCCCTAAATGAGTGAGCTAAACTCA 831
Db	129	ACATACGAGCCGGAGCATAAAGTGTAAAGCCTCGGGTGCCCTAAATGAGTGAGCTAAACTCA 188
QY	832	CATTAAATTCGGTTGCGCTCAGCTGCCGCGTTTCCAGTTCGGGAAACCTGTCGTCCAGCTGC 891
Db	189	CATTAAATTCGGTTGCGCTCAGCTGCCGCGTTTCCAGTTCGGGAAACCTGTCGTCCAGCTGC 248
QY	892	ATTAAATGAATCGGCCCAACCGCGGGGAGAGCGCGTTTTCGGTATTGGGCGCTCTTCCCGTTT 951
Db	249	ATTAAATGAATCGGCCCAACCGCGGGGAGAGCGCGTTTTCGGTATTGGGCGCTCTTCCCGTTT 308
QY	952	CCTTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTCGCGCGAGCGGTA 1001

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:10:26 ; Search time 3961.99 Seconds
(without alignments)
10335.846 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 cgcgtggaggatcatccagcc.....gccatcagatcccttggcggc 1001

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.in.*
- 32: em.htg.inv.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.vrt.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	1001	100.0	1300	1	ISTN5X	V00618 Transposon
C 2	1001	100.0	1300	6	I01012	I01012 Sequence 4
C 3	1001	100.0	1500	6	E02455	E02455 DNA encodin
C 4	1001	100.0	2220	6	AR067785	AR067785 Sequence
C 5	1001	100.0	2320	12	SYNKANRA	M84115 Cloning vec
C 6	1001	100.0	2356	1	ECU32991	U32991 Escherichia
C 7	1001	100.0	3267	12	AY048743	AY048743 Template
C 8	1001	100.0	3740	6	A00782	A00782 Plasmid pDM
C 9	1001	100.0	3740	6	A02231	A02297 Plasmid pDM
C 10	1001	100.0	3740	6	A02297	A07054 Artificial
C 11	1001	100.0	3740	6	A07054	A0362 Artificial
C 12	1001	100.0	3740	6	A10362	A14595 Synthetic n
C 13	1001	100.0	3740	6	A14595	A19608 Artificial
C 14	1001	100.0	3740	6	A19608	AR096518 Sequence
C 15	1001	100.0	3740	6	A30452	AR096518 Sequence
C 16	1001	100.0	3974	6	AR096518	AR098395 Sequence
C 17	1001	100.0	3974	6	AR098395	AR099244 Sequence
C 18	1001	100.0	3974	6	AR099244	AR172638 Sequence
C 19	1001	100.0	3974	6	AR172638	AR207786 Sequence
C 20	1001	100.0	3974	6	AR207786	AR218546 Sequence
C 21	1001	100.0	3974	6	AR218546	AR253199 Sequence
C 22	1001	100.0	3974	6	AR253199	AR256379 Sequence
C 23	1001	100.0	3974	6	AR256379	AR267137 Sequence
C 24	1001	100.0	3974	6	AR267137	BD022735 Mammary c
C 25	1001	100.0	3974	6	BD022735	BD033415 Anticirc
C 26	1001	100.0	3974	6	BD033415	BD129548 Polynuc
C 27	1001	100.0	3974	6	BD129548	BD131214 Apoptosis
C 28	1001	100.0	3974	6	BD131214	BD134158 Vascular
C 29	1001	100.0	3974	6	BD134158	BD134899 Apoptosis
C 30	1001	100.0	3974	6	BD134899	BD136632 Therapeut
C 31	1001	100.0	3974	6	BD136632	AR134670 Sequence
C 32	1001	100.0	3984	6	AR134670	AR219194 Sequence
C 33	1001	100.0	3984	6	AR219194	AX589681 Sequence
C 34	1001	100.0	3997	6	AX589681	AR134669 Sequence
C 35	1001	100.0	4208	6	AR134669	AR219193 Sequence
C 36	1001	100.0	4208	6	AR219193	AR256415 Sequence
C 37	1001	100.0	4208	6	AR256415	BD005772 Therapeut
C 38	1001	100.0	4208	6	BD005772	AR207828 Sequence
C 39	1001	100.0	4256	6	AR207828	AR134671 Sequence
C 40	1001	100.0	4277	6	AR134671	AR219195 Sequence
C 41	1001	100.0	4277	6	AR219195	A92666 Sequence 2
C 42	1001	100.0	4328	6	A92666	AR137192 Sequence
C 43	1001	100.0	4328	6	AR137192	AR158346 Sequence
C 44	1001	100.0	4328	6	AR158346	AR241208 Sequence
C 45	1001	100.0	4328	6	AR241208	

ALIGNMENTS

RESULT 1
ISTN5X/c
LOCUS
DEFINITION
Transposon Tn5 fragment encoding neomycin and kanamycin resistance (neo) and a fragment of the reading frame of a further protein.
V00618 J01834
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia

ISTN5X 1300 bp DNA linear BCT 02-SBP-1999
Transposon Tn5 fragment encoding neomycin and kanamycin resistance (neo) and a fragment of the reading frame of a further protein.

REFERENCE
AUTHORS
TITLE
Beck, E., Ludwig, G., Auerwald, E.A., Reiss, B. and Schaller, H.
Nucleotide sequence and exact localization of the neomycin

Phosphotransferase gene from transposon Tn5

JOURNAL
Gene 19 (3), 327-336 (1982)
MEDLINE
83106478
PUBMED
6295884

FEATURES

source

1. .1300
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
151_945
/codon_start=1
/transl_table=11
/product="neomycin phosphotransferase"
/protein_id="CAA23892.1"
/db_xref="GI:43749"
/db_xref="SWISS-PROT:P00552"
/translation="MTEDGLHAGSPAAWVERLFGVDNAQOTIGSDAAVRLSAQGR
PVLFWKTLGSLNELQDEAPRLSWLATGVPCAALVIVTEAGRDWLLGLEVPQDL
LSHLAPAKYSLIMADNRRLTLPALCPFDQAKHRIERARKMERGLVDQDLDE
EQGLAPALFARLARMPPDEDLVVTHGDACLPMIVENGRFSGFDICGLGVADRY
QDIALATRDIAELGGEWADRFVLVYGAAPDSQRIAFYRLIDEPF"
942..>1300

CDS

1. .1300
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
151_945
/codon_start=1
/transl_table=11
/product="neomycin phosphotransferase"
/protein_id="CAA23892.1"
/db_xref="GI:43749"
/db_xref="SWISS-PROT:P00552"
/translation="MTEDGLHAGSPAAWVERLFGVDNAQOTIGSDAAVRLSAQGR
PVLFWKTLGSLNELQDEAPRLSWLATGVPCAALVIVTEAGRDWLLGLEVPQDL
LSHLAPAKYSLIMADNRRLTLPALCPFDQAKHRIERARKMERGLVDQDLDE
EQGLAPALFARLARMPPDEDLVVTHGDACLPMIVENGRFSGFDICGLGVADRY
QDIALATRDIAELGGEWADRFVLVYGAAPDSQRIAFYRLIDEPF"
942..>1300

CDS

1. .1300
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
151_945
/codon_start=1
/transl_table=11
/product="neomycin phosphotransferase"
/protein_id="CAA23892.1"
/db_xref="GI:43749"
/db_xref="SWISS-PROT:P00552"
/translation="MTEDGLHAGSPAAWVERLFGVDNAQOTIGSDAAVRLSAQGR
PVLFWKTLGSLNELQDEAPRLSWLATGVPCAALVIVTEAGRDWLLGLEVPQDL
LSHLAPAKYSLIMADNRRLTLPALCPFDQAKHRIERARKMERGLVDQDLDE
EQGLAPALFARLARMPPDEDLVVTHGDACLPMIVENGRFSGFDICGLGVADRY
QDIALATRDIAELGGEWADRFVLVYGAAPDSQRIAFYRLIDEPF"
942..>1300

BASE COUNT 245 a 374 c 408 g 273 t

ORIGIN

Query Match 100.0%; Score 1001; DB 1; Length 1300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 60
DB 1084 CGCTGAGGATATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 1025

QY 61 AAGCGCGGTGGAATCGAAATCTCGTGATGGCAGGTGGCGCTGCGTTCATG 120
DB 1024 AAGCGCGGTGGAATCGAAATCTCGTGATGGCAGGTGGCGCTGCGTTCATG 965

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGTGCGCT 180
DB 964 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGTGCGCT 905

QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCACAGGAGCGGTGAGCCCAATTCGCGCCAA 240
DB 904 GCGAATCGGAGCGCGGATACCGTAAAGCACAGGAGCGGTGAGCCCAATTCGCGCCAA 845

QY 241 GCTCTTCAGCAATATCAGCGGTAGCAAGCGTATGTCCTGATAGCGGTCCGCCACACCA 300
DB 844 GCTCTTCAGCAATATCAGCGGTAGCAAGCGTATGTCCTGATAGCGGTCCGCCACACCA 785

QY 301 GCGGCGCACAGTGCATGAATCCAGAAAGCGGCAATTTCCACCATGATATTCGCGCAAGC 360
DB 784 GCGGCGCACAGTGCATGAATCCAGAAAGCGGCAATTTCCACCATGATATTCGCGCAAGC 725

QY 361 AGGCATCGCCATGGGTCAAGAGATCTCGCGCTCGGCGATGCGCGCTTTCAGCGCTGG 420
DB 724 AGGCATCGCCATGGGTCAAGAGATCTCGCGCTCGGCGATGCGCGCTTTCAGCGCTGG 665

QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTTCATCGACAA 480
DB 664 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTTCATCGACAA 605

QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
DB 604 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 545

QY 541 GGCAGGTAGCCGGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATACCT 600
DB 544 GGCAGGTAGCCGGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATACCT 485

QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATTCCTGCCCCGGCACTTCGCCCAATAGCA 660
DB 484 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATTCCTGCCCCGGCACTTCGCCCAATAGCA 425

QY 661 GCCAGTCCCTTCCCGCTTCAGTGCACACGTGCGAGCACAGTGGCGAGAAAGCCCGCTCG 720
DB 424 GCCAGTCCCTTCCCGCTTCAGTGCACACGTGCGAGCACAGTGGCGAGAAAGCCCGCTCG 365

QY 721 TGCGCAGCCAGATAGCGCGCTCGCTTCCTGCAAGTTCATTCAGGCGCACCGGACAGGT 780
DB 364 TGCGCAGCCAGATAGCGCGCTCGCTTCCTGCAAGTTCATTCAGGCGCACCGGACAGGT 305

QY 781 CGCTCTTGACAAAAGAACCGGCGCGCTCGCTGCACAGCGGAAACGAGCGGCGCATCAG 840
DB 304 CGCTCTTGACAAAAGAACCGGCGCGCTCGCTGCACAGCGGAAACGAGCGGCGCATCAG 245

QY 841 AGCAGCGGATGCTGTGTGTCGCCAGTCTATAGCGCGAATAGCCTCTCCACCCAAAGCGCGCG 900
DB 244 AGCAGCGGATGCTGTGTGTCGCCAGTCTATAGCGCGAATAGCCTCTCCACCCAAAGCGCGCG 185

QY 901 GAGAACCTGCTGCATTCATCTTGTTCATATGCGAAACGATCTCTCTCTCTCTCTCT 960
DB 184 GAGAACCTGCTGCATTCATCTTGTTCATATGCGAAACGATCTCTCTCTCTCTCTCT 125

QY 961 GATCAGATCTTGTATCCCTGCGGCATCAGATCTTCTGCGCGC 1001
DB 124 GATCAGATCTTGTATCCCTGCGGCATCAGATCTTCTGCGCGC 84

RESULT 2
LOCUS I01012/c 1300 bp ss-DNA linear PAT 21-MAY-1993
DEFINITION Sequence 4 from Patent US 4782022.
ACCESSION I01012
VERSION I01012.1 GI:269592
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1300)
AUTHORS Puhler, A., Reilaender, H. and Weber, G.
TITLE Nitrogen fixation regulator genes.
JOURNAL Patent: US 4782022-A 4 01-NOV-1988;
Lubrizol Genetics, Inc.; Wickliffe, OH
FEATURES
source
1. .1300
/organism="unknown"

BASE COUNT 245 a 374 c 408 g 273 t
ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 1300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 60
DB 1084 CGCTGAGGATATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 1025

QY 61 AAGCGCGGTGGAATCGAAATCTCGTGATGGCAGGTGGCGCTGCGTTCATG 120
DB 1024 AAGCGCGGTGGAATCGAAATCTCGTGATGGCAGGTGGCGCTGCGTTCATG 965

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGTGCGCT 180
DB 964 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGTGCGCT 905

QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCACAGGAGCGGTGAGCCCAATTCGCGCCAA 240

Db 904 GCGATCGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCGGCCAA 845
QY 241 GCTCTTACGCAATATACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 300
Db 844 GCTCTTACGCAATATACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 785
QY 301 GCCGCCACAGTCGATGATCCAGAAAGCGCCATTTCCACCATGATATTCGGCAGC 360
Db 784 GCCGCCACAGTCGATGATCCAGAAAGCGCCATTTCCACCATGATATTCGGCAGC 725
QY 361 AGGCATCGCATGGGTACAGAGATCTCGCCGTGCGGATCGCGCCCTTGAGCCTGG 420
Db 724 AGGCATCGCATGGGTACAGAGATCTCGCCGTGCGGATCGCGCCCTTGAGCCTGG 665
QY 421 CGAACAGTTCGGTCGGCGAGCCCTGATGCTTCTGTCCAGATCATCTGTATCGACAA 480
Db 664 CGAACAGTTCGGTCGGCGAGCCCTGATGCTTCTGTCCAGATCATCTGTATCGACAA 605
QY 481 GACCCGCTTCCATCCGAGTACGTGCTCGCTCGATCGCATGTTTCGTTGGTGGTGAATG 540
Db 604 GACCCGCTTCCATCCGAGTACGTGCTCGCTCGATCGCATGTTTCGTTGGTGGTGAATG 545
QY 541 GCGAGGTACCGGATCAAGGTATGACGCGCGCATTTGCATCAGCCCATGATGATCTT 600
Db 544 GCGAGGTACCGGATCAAGGTATGACGCGCGCATTTGCATCAGCCCATGATGATCTT 485
QY 601 TCTCGGAGGACGAGGTGAGTACAGAGATCTCGCCCGGACCTTCGCCCATAGCA 660
Db 484 TCTCGGAGGACGAGGTGAGTACAGAGATCTCGCCCGGACCTTCGCCCATAGCA 425
QY 661 GCCAGTCCCTTCCGCTTCAGTGACAGCTGACAGCATCTCGCAAGGACGCGCGTCG 720
Db 424 GCCAGTCCCTTCCGCTTCAGTGACAGCTGACAGCATCTCGCAAGGACGCGCGTCG 365
QY 721 TGGCAGCAGATAGCGCGTCTGCTCTGTCAGATTCATTCAGGACCGGACAGGT 780
Db 364 TGGCAGCAGCAGATAGCGCGTCTGCTCTGTCAGATTCATTCAGGACCGGACAGGT 305
QY 781 CGGTCTTGACAAAAGAACCGGCGCCCTCGCTGACAGCGGAAACAGCGCGCATCAG 840
Db 304 CGGTCTTGACAAAAGAACCGGCGCCCTCGCTGACAGCGGAAACAGCGCGCATCAG 245
QY 841 AGCAGCGATGCTGTGTGGCCAGTATAGCGGATAGCCTCTCCACCCAGCGCGCG 900
Db 244 AGCAGCGATGCTGTGTGGCCAGTATAGCGGATAGCCTCTCCACCCAGCGCGCG 185
QY 901 GAGAACCTCGGTGCAATCCATCTGTTCAATCATCGGAAAGATCTCTCTCTCTT 960
Db 184 GAGAACCTCGGTGCAATCCATCTGTTCAATCATCGGAAAGATCTCTCTCTCTT 125
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCTCTTGGCGG 1001
Db 124 GATCAGATCTTGATCCCTCGCCATCAGATCTTGGCGG 84

RESULT 3
E02455/c
LOCUS E02455 1600 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding aminoglycoside 3'-phosphotransferaseII (APH).
ACCESSION E02455
VERSION E02455 1 GI:2170686
KEYWORDS JP 1990150282-A/1.
SOURCE Transposon Tn5
ORGANISM transposons.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Yoshida,N., Kikuchi,N., Shin,M. and Teraoka,H.
TITLE MODIFIED HUMAN PSTI
JOURNAL Patent: JP 1990150282-A 1 08-JUN-1990;
SHIONOGI & CO LTD
OS Transposon Tn5
FN JP 1990150282-A/1
PD 08-JUN-1990

PF 11-OCT-1988 JP 1988255580
PR 19-JUL-1988 JP 88P 181316
PI YOSHIDA NOBUO, KIKUCHI NORIHISA, SHIN MASARU, TERAOKA HIROSHI
PC C12N15/12,C07K7/10//C07H21/04,C12P21/02,(C12P21/02,C12R1:19),
C07K9:00;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pNEO;
Key Location/Qualifiers
FT 5'UTR 1..450
FT CDS 451..1245
FT /product='APH'
FT mat_peptide 451..1242
FT /product='APH'
FT 3'UTR 1246..1600.
FT Location/Qualifiers
1..1600
/organism='Transposon Tn5'
/mol_type='genomic DNA'
/db_xref='taxon:2411'
BASE COUNT 324 a 442 c 508 g 326 t
ORIGIN
Query Match 100.0%; Score 1001; DB 6; Length 1600;
Best Local Similarity 100.0%; Freq. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTGAGGATCATCCAGCGCGGTCCCGAAAAAGATTCCGAAGCCCAACCTTTTCATAG 60
Db 1384 CGTGTGAGGATCATCCAGCGCGGTCCCGAAAAAGATTCCGAAGCCCAACCTTTTCATAG 1325
QY 61 AAGCGCGCGTGAATCGAAATCTGTGATGCGAGGTTGGGCGTGTCTTGGTGGTCAAT 120
Db 1324 AAGCGCGCGTGAATCGAAATCTGTGATGCGAGGTTGGGCGTGTCTTGGTGGTCAAT 1265
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGACTCGTCAAGAGGCGATAGAGGCGATCGCT 180
Db 1264 TCGAACCCAGAGTCCCGCTCAGAAAGACTCGTCAAGAGGCGATAGAGGCGATCGCT 1205
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCGCCAA 240
Db 1204 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCGCCAA 1145
QY 241 GCTCTTACGCAATATACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 300
Db 1144 GCTCTTACGCAATATACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 1085
QY 301 GCGGCCACAGTCGATGATCCAGAAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 1084 GCGGCCACAGTCGATGATCCAGAAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 1025
QY 361 AGGCATCGCATGGGTACAGAGATCTCGCCGTGCGGATCGCGGCTTGAGCCTGG 420
Db 1024 AGGCATCGCATGGGTACAGAGATCTCGCCGTGCGGATCGCGGCTTGAGCCTGG 965
QY 421 CGAACAGTTCGGTCGGCGAGCCCTGATGCTTCTGTCCAGATCATCTGTATCGACAA 480
Db 964 CGAACAGTTCGGTCGGCGAGCCCTGATGCTTCTGTCCAGATCATCTGTATCGACAA 905
QY 481 GACCGGCTTCCATTCGAGTACGTGCTCGATCGATGTTTCGTTGGTGGTGAATG 540
Db 904 GACCGGCTTCCATTCGAGTACGTGCTCGATCGATGTTTCGTTGGTGGTGAATG 845
QY 541 GCGAGGTACCGGATCAAGCGGTATGACGCGCGGATTCATCAGCATGATGATCTT 600
Db 844 GCGAGGTACCGGATCAAGCGGTATGACGCGCGGATTCATCAGCATGATGATCTT 785
QY 601 TCTCGCAGGACGAGGTGAGATGACAGAGATCTCTGCCCGGCACTTCGCCCAATAGCA 660
Db 784 TCTCGCAGGACGAGGTGAGATGACAGAGATCTCTGCCCGGCACTTCGCCCAATAGCA 725

QY 661 GCCAGTCCCTTCCGCTTCACTGAGCAACAGCTCGAGCAGAGTGGCGAAGCAAGCCGCTCG 720
Db 724 GCCAGTCCCTTCCGCTTCACTGAGCAACAGCTCGAGCAGAGTGGCGAAGCAAGCCGCTCG 665
QY 721 TGCCAGCCAGATAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 664 TGCCAGCCAGATAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
QY 781 CGGTCTTGACAAAAGAACCGGCGCCCTGACGCTGACGCTGACGCTGACGCTGACGCTGAC 840
Db 604 CGGTCTTGACAAAAGAACCGGCGCCCTGACGCTGACGCTGACGCTGACGCTGACGCTGAC 545
QY 841 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 544 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
QY 901 GAGAACCTGGTGAATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 484 GAGAACCTGGTGAATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
QY 961 GATCAGATCTTGATCCCTCGGCAATAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 424 GATCAGATCTTGATCCCTCGGCAATAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 384

RESULT 4

LOCUS AR067785 2220 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5851808.
ACCESSION AR067785
VERSION AR067785.1 GI:5999007
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2220)
AUTHORS Elledge, S.J. and Liu, Q.
TITLE Rapid subcloning using site-specific recombination
JOURNAL Patent: US 5851808-A 1 22-DEC-1998;
FEATURES Location/Qualifiers
source 1..2220
/organism="unknown"
BASE COUNT 506 a 610 c 581 g 523 t
ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 2220;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATCCGAAAGCCCAACCTTTCATAG 60
Db 932 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATCCGAAAGCCCAACCTTTCATAG 991
QY 61 AAGCGCGGTGGAAATCGAATCTGCTGATGGCAGGTGGCGGTGCTGCTGCTGCTGCTGCT 120
Db 992 AAGCGCGGTGGAAATCGAATCTGCTGATGGCAGGTGGCGGTGCTGCTGCTGCTGCTGCT 1051
QY 121 TCGAACCCACAGTCCCGCTCAGAGAACCTGTCAGAGGCGATAGAGGCGATCGCT 180
Db 1052 TCGAACCCACAGTCCCGCTCAGAGAACCTGTCAGAGGCGATAGAGGCGATCGCT 1111
QY 181 CGCAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGAGCGGCTGAGCGGCTTCCGCCCAA 240
Db 1112 CGCAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGAGCGGCTGAGCGGCTTCCGCCCAA 1171
QY 241 GCTCTTCAGCATATACGGGTAGCCAAACGCTATGCTGCTGATAGCGGTGCGCCACACCA 300
Db 1172 GCTCTTCAGCATATACGGGTAGCCAAACGCTATGCTGCTGATAGCGGTGCGCCACACCA 1231
QY 301 GCCGCCACAGTCCATGAATCCAGAAAGCGGCGCTTTCCACCATGATATTCGCGAAGC 360
Db 1232 GCCGCCACAGTCCATGAATCCAGAAAGCGGCGCTTTCCACCATGATATTCGCGAAGC 1291

QY 361 AGCATCGCATGGGTCAAGCAGAGATCTTCGCGCTCGGCGATCGCGCTTGAGCCTGG 420
Db 1292 AGCATCGCATGGGTCAAGCAGAGATCTTCGCGCTCGGCGATCGCGCTTGAGCCTGG 1351
QY 421 CGAACAGTTCGGCTTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 1352 CGAACAGTTCGGCTTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 1411
QY 481 GACCGGCTTCATCCGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 1412 GACCGGCTTCATCCGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1471
QY 541 GCGAGGTAGCGGATCAAGCGTATGACGCGCGCGCATTTGATCAGCCATCATGATGATACTT 600
Db 1472 GCGAGGTAGCGGATCAAGCGTATGACGCGCGCGCATTTGATCAGCCATCATGATGATACTT 1531
QY 601 TCTCGCAGAGCAAGGTGAGTACAGAGAGATCTTCGCGCGGCACTTCGCCCAATAGCA 660
Db 1532 TCTCGCAGAGCAAGGTGAGTACAGAGAGATCTTCGCGCGGCACTTCGCCCAATAGCA 1591
QY 661 GCCAGTCCCTTCCGCTTCACTGACAACTGACAGCAGCTGCGCAAGGAACGCCCGCTCG 720
Db 1592 GCCAGTCCCTTCCGCTTCACTGACAACTGACAGCAGCTGCGCAAGGAACGCCCGCTCG 1651
QY 721 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1652 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
QY 781 CGGTCTTGACAAAAGAACCGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1712 CGGTCTTGACAAAAGAACCGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1771
QY 841 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 1772 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
QY 901 GAGAACCTGGTGAATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1832 GAGAACCTGGTGAATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
QY 961 GATCAGATCTTGATCCCTCGGCAATAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 1892 GATCAGATCTTGATCCCTCGGCAATAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932

RESULT 5

LOCUS SYNKANRA
DEFINITION Cloning vector pHS96 kanamycin resistance protein (neo) gene, putative cds.
ACCESSION M84115
VERSION M84115.1 GI:146531
KEYWORDS kanamycin resistance, unidentified cloning vector
SOURCE unidentified cloning vector
ORGANISM artificial sequences; vectors.
REFERENCE 1 (sites)
AUTHORS Seifert, H.S., So, M. and Heffron, F.
TITLE Shuttle mutagenesis: a method of introducing transposons into transformable organisms
JOURNAL Genet. Eng. Prin. Methods 8, 327-336 (1986)
REFERENCE 2 (bases 1 to 2320)
AUTHORS Seifert, H.S., Chen, E.Y., So, M. and Heffron, F.
TITLE Shuttle mutagenesis: a method of transposon mutagenesis for Saccharomyces cerevisiae
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (3), 735-739 (1986)
MEDLINE 86121014
PUBMED 3003748
COMMENT Original source text: Cloning vector (sub-species Cloning vector pHS96) DNA.
FEATURES Cloning vector pHS96 is used in the system of shuttle mutagenesis. Location/Qualifiers

```

source
1. .2320
/organism="unidentified cloning vector"
/mol_type="genomic DNA"
/sub_species="Cloning vector pHS86"
/db_xref="taxon:45196"
complement(1232..2026)
/gene="neo"
complement(1232..2026)
/misc_feature
/note="putative cds; putative"
/function="kanamycin resistance gene from Tn5"
1442
/standard_name="oriM13"
/note="M13 origin in derivative pHS7; single stranded DNA
replication origin; a 191 bp fragment from pS1 (Johnston
& Ray, JMB 177:685-700, 1984) was inserted at this
position to make the derivative pHS7"
/evidence="experimental"
1660
/rep_origin
/note="putative"
/direction=left
509 a 698 c 619 g 494 t
BASE COUNT
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2320;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCACCGCGCTCCGGAAGAGATTCGGAAGCCACCTTTCATAG 60
DB 155 CGCTGAGGATCATCCACCGCGCTCCGGAAGAGATTCGGAAGCCACCTTTCATAG 214

QY 61 AAGGGGGGGTGGAAATCTGTCATGCGAGGTTGGCGTCGCTTGGTCGGTCATT 120
DB 215 AAGGGGGGGTGGAAATCTGTCATGCGAGGTTGGCGTCGCTTGGTCGGTCATT 274

QY 121 TCGAACCCAGAGTCCCTCAGAGAACTGTCAGAGCGGATAGAGCGATCGGCT 180
DB 275 TCGAACCCAGAGTCCCTCAGAGAACTGTCAGAGCGGATAGAGCGATCGGCT 334

QY 181 GCGAATCGGGCGGATGATCCAGAAAGCGGAGCGGTCAGCCCATTCGCCGCCAA 240
DB 335 GCGAATCGGGCGGATGATCCAGAAAGCGGAGCGGTCAGCCCATTCGCCGCCAA 394

QY 241 GCTCTTCAGCATATCAGGTCAGGATGATGCTGTCAGAGCGGTCGCGCACACCA 300
DB 395 GCTCTTCAGCATATCAGGTCAGGATGATGCTGTCAGAGCGGTCGCGCACACCA 454

QY 301 GCGGCGCAGTCGATGAATCCAGAAAGCGGCGCCATTTCCACCATGATTCGCAAGC 360
DB 455 GCGGCGCAGTCGATGAATCCAGAAAGCGGCGCCATTTCCACCATGATTCGCAAGC 514

QY 361 AGGCATCGCCATGGTTCAGGACGATCTCGCGTCGGGATCGCGGCTTTCAGCGTGG 420
DB 515 AGGCATCGCCATGGTTCAGGACGATCTCGCGTCGGGATCGCGGCTTTCAGCGTGG 574

QY 421 CGAACAGTTTCGGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATGACAA 480
DB 575 CGAACAGTTTCGGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATGACAA 634

QY 481 GACCGGCTTCATTCAGTATGATGCTCGCTGATGCGATGTTTCGCTTGTGTGCAATG 540
DB 635 GACCGGCTTCATTCAGTATGATGCTCGCTGATGCGATGTTTCGCTTGTGTGCAATG 694

QY 541 GGCAGGTAGCGGATCAAGCGTATGTCAGCGCCGCGCATTTGCATCAGCCATGATGATCT 600
DB 695 GGCAGGTAGCGGATCAAGCGTATGTCAGCGCCGCGCATTTGCATCAGCCATGATGATCT 754

QY 601 TCTCGCGGAGGACAGGTGAGTGCAGAGATCTCTGCGCGGACCTTCCGCAATAGCA 660
DB 755 TCTCGCGGAGGACAGGTGAGTGCAGAGATCTCTGCGCGGACCTTCCGCAATAGCA 814

QY 661 GCGAGTCCCTTCCCGCTTCAGTGCACACGTCGAGCAGAGTCGCGCAAGAACCGCGTGC 720

```

815 GCCAGTCCCTTCCCGCTTCAGTGCACACGTCGAGCAGAGTCGCGCAAGAACGCCCGCTCG 874

721 TGGCCAGCCACGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

875 TGGCCAGCCACGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934

781 CGGCTTTGACAAAAGAAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

935 CGGCTTTGACAAAAGAAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994

841 AGCAGCCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

995 AGCAGCCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054

901 GAGAACCTCGGTCGAATCCATCTTGTTCATCATGCGAAACGATCTCATCTGCTGCTT 960

1055 GAGAACCTCGGTCGAATCCATCTTGTTCATCATGCGAAACGATCTCATCTGCTGCTT 1114

961 GATCAGATCTTGTATCCCTCGCGCATCAGATCTTGGCGGC 1001

1115 GATCAGATCTTGTATCCCTCGCGCATCAGATCTTGGCGGC 1155

ECU32991/c

LOCUS ECU32991 2356 bp DNA linear BCT 23-JAN-1996

DEFINITION Escherichia coli mini-Tn5 kanamycin transposon.

ACCESSION U32991

VERSION U32991.1 GI:1163175

KEYWORDS

SOURCE Escherichia coli

ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 2356)

AUTHORS de Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N.

TITLE Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria

JOURNAL J. Bacteriol. 172 (11), 6568-6572 (1990)

MEDLINE 91035272

PUBMED 2172217

REFERENCE 2 (bases 1 to 2356)

AUTHORS Pinyon, R.A. and Thomas, C.J.

TITLE DNA sequence of a mini-Tn5 transposon (mini-Tn5 Km)

JOURNAL Unpublished

REFERENCE 3 (sites)

AUTHORS Herrero, M., de Lorenzo, V. and Timmis, K.N.

TITLE Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria

JOURNAL J. Bacteriol. 172 (11), 6557-6567 (1990)

MEDLINE 91035271

PUBMED 2172216

REFERENCE 4 (sites)

AUTHORS Prenk, P. and Krisch, H.M.

TITLE In vitro insertional mutagenesis with a selectable DNA fragment

JOURNAL Gene 29 (3), 303-313 (1984)

MEDLINE 85028451

PUBMED 6237955

REFERENCE 5 (sites)

AUTHORS Beck, S., Ludwig, G., Auerwald, E.A., Reiss, B. and Schaller, H.

TITLE Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5

JOURNAL Gene 19 (3), 327-336 (1982)

MEDLINE 83106478

PUBMED 6295884

REFERENCE 6 (bases 1 to 2356)

AUTHORS Pinyon, R.A.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-1995) Rebecca A. Pinyon, Department of Microbiology and Immunology, University of Adelaide, Frome Road,

Adelaide, S.A. 5001, Australia
 On Jan 23, 1996 this sequence version replaced gi:1000122.
 COMMENT Location/Qualifiers
 FEATURES
 source
 1. .2356
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /strain="SM10 (lambda pir)"
 /db_xref="taxon:562"
 /plasmid="mini-Tn5 Km"
 /transposon="mini-Tn5 Km"
 repeat_region
 1. .19
 /note="Tn5 I end functional in transposition"
 /citations=[4]
 /rpt_type=inverted
 857..1651
 CDS
 /functions="neomycin and kanamycin resistance"
 /codon_start=11
 /transl_table=11
 /product="neomycin phosphotransferase"
 /protein_id="AA85506.1"
 /db_xref="GI:1000123"
 /translations="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAVRLSAQGR
 LSHLAPAEKVSIMADAMRLHLTDPATCFDFOAKHRIERARTRMEAGLVDDDLDE
 EHOGLAPAEIAPLAKAMPDGEDLVTHGDACLEIMVENGEPSPIDGRLGVADRY
 QDALATRDIAELGSEWADRFVLVYGLIAPDSQRIAYRLLDEFF"
 repeat_region
 2338..2356
 /note="Tn5 O end functional in transposition"
 /citations=[4]
 /rpt_type=inverted
 601 c 548 t
 BASE COUNT 532 a 601 c 675 g 548 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 1; Length 2356;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCCGAAGCCCAACCTTTCATAG 60
 DB 1790 CGCTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCCGAAGCCCAACCTTTCATAG 1731

QY 61 AAGCGGGTGGGATCGAATCTCGTGATGGCAGCTGGGGTGGCTGGTGGCTGATT 120
 DB 1730 AAGCGGGTGGGATCGAATCTCGTGATGGCAGCTGGGGTGGCTGGTGGCTGATT 1671

QY 121 TCGAACCCAGATCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAGCGATGCGCT 180
 DB 1670 TCGAACCCAGATCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAGCGATGCGCT 1611

QY 181 CGGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTACGCCATTCGCCGCCAA 240
 DB 1610 CGGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTACGCCATTCGCCGCCAA 1551

QY 241 GCTCTTCAGCAATATACGGGTAGCCAAACCTATGCTGATAGCGGTCCGCCAACCCCA 300
 DB 1550 GCTCTTCAGCAATATACGGGTAGCCAAACCTATGCTGATAGCGGTCCGCCAACCCCA 1491

QY 301 GCCGCCACAGTCGATCAATCCAGAAAGCGGCATTTCCACCATGATATTCGCGAAGC 360
 DB 1490 GCCGCCACAGTCGATCAATCCAGAAAGCGGCATTTCCACCATGATATTCGCGAAGC 1431

QY 361 AGGCATGCCATGGGTACGACGAGATTCCTCGCGTCCGCCATGCGCGCTTTCAGCGCTG 420
 DB 1430 AGGCATGCCATGGGTACGACGAGATTCCTCGCGTCCGCCATGCGCGCTTTCAGCGCTG 1371

QY 421 CGAACAGTTCGGTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTCATCGACAA 480
 DB 1370 CGAACAGTTCGGTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTCATCGACAA 1311

QY 481 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
 DB 1310 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 1251

QY 541 GCGAGTAGCCGGATCAAGCGTATGCGAGCCCGCGGANTTCATGAGCATGATGATGACTT 600
 DB 1250 GCGAGTAGCCGGATCAAGCGTATGCGAGCCCGCGGANTTCATGAGCATGATGATGACTT 1191

QY 601 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGCGCACTTCGCCCAATAGCA 660
 DB 1190 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGCGCACTTCGCCCAATAGCA 1131

QY 661 GCGAGTCCCTTCGCGTTCAGTGACAGCTGAGCAGAGCTGCGCAAGGAGCGCCCTCG 720
 DB 1130 GCGAGTCCCTTCGCGTTCAGTGACAGCTGAGCAGAGCTGCGCAAGGAGCGCCCTCG 1071

QY 721 TGCGCCAGCCAGATAGCCGCTGCTGCTCTCTGAGTTTCAATCAGGCGCACCGACAGGT 780
 DB 1070 TGCGCCAGCCAGATAGCCGCTGCTGCTCTCTGAGTTTCAATCAGGCGCACCGACAGGT 1011

QY 781 CGGTCTTGCAAAAGAAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 1010 CGGTCTTGCAAAAGAAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951

QY 841 AGCAGCCGATTTCTGTTGTGCCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCGCG 900
 DB 950 AGCAGCCGATTTCTGTTGTGCCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCGCG 891

QY 901 GAGACCTCGTGCATCACTTGTTCATCATGCGAAGCGATCTCATCTCTCTCTT 960
 DB 890 GAGACCTCGTGCATCACTTGTTCATCATGCGAAGCGATCTCATCTCTCTCTT 831

QY 961 GATCAGATCTTGATCCCTGCGCCATCAGATCTTGGCGGC 1001
 DB 830 GATCAGATCTTGATCCCTGCGCCATCAGATCTTGGCGGC 790

RESULT 7

AY048743/c
 LOCUS AY048743 3267 bp DNA circular SYN 11-SEP-2001
 DEFINITION Template plasmid pKD4, complete sequence.
 ACCESSION AY048743
 VERSION AY048743.1 GI:15554332
 KEYWORDS
 SOURCE Template plasmid pKD4
 ORGANISM
 Datsenko, K.A. and Wanner, B.L.
 One-step inactivation of chromosomal genes in Escherichia coli K-12
 using PCR products
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6640-6645 (2000)
 MEDLINE 20300954
 PUBMED 10829079
 REFERENCE 2 (bases 1 to 3267)
 Datsenko, K.A. and Wanner, B.L.
 Direct Submission
 TITLE Submitted (29-JUL-2001) Biological Sciences, Purdue University,
 Lilly Hall of Life Sciences, West Lafayette, IN 47907, USA
 JOURNAL
 FEATURES
 Location/Qualifiers
 1. .3267
 /organism="Template plasmid pKD4"
 /mol_type="genomic DNA"
 /db_xref="taxon:170492"
 31..50
 /note="priming site 1"
 32..51
 /note="priming site 3"
 complement(51..98)
 /note="natural FRT site"
 459..1253
 /note="kanamycin resistance"
 /codon_start=11
 /transl_table=11
 /product="Tn5 neomycin phosphotransferase"
 /protein_id="AA02035.1"
 /db_xref="GI:15554333"

misc_feature
 31..50
 /note="priming site 1"
 misc_feature
 32..51
 /note="priming site 3"
 complement(51..98)
 /note="natural FRT site"
 459..1253
 /note="kanamycin resistance"
 /codon_start=11
 /transl_table=11
 /product="Tn5 neomycin phosphotransferase"
 /protein_id="AA02035.1"
 /db_xref="GI:15554333"


```

/misc_feature
529.548
/notes="common priming site k1"
/misc_feature
620.639
/notes="common priming site k2"
/misc_feature
1090.1109
/notes="common priming site kt"
/misc_feature
1444.11489
/notes="FRT site"
1507.1526
/notes="priming site 2"
1523.1909
/notes="R6K gamma replication origin"
complement (1979, 2839)
/notes="bla; ampicillin resistance"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAL02036.1"
/db_xref="GI:15554334"
/translation="MSIQHFRVALIPFPAAFLPVFAHPETLVKVDAAEDQLGARVGY
IELDLNSGKILDSFEPERPPNWSFKVLGCVLSDIDAGEOEOLGERIRHYSDNIVE
YSPVTEKHLTDGMYRELCSAAITMSDNTAANLLITIGGPRELTAFLHNNQDVIKRL
LRWPELNAIDENRDTIPVMATIRKLLTGELTLASRQLLIQDMEADKVVAGEL
LRSLPAGWFIADKSGAGERSGIIAALGPDGKPSRIIVYITGTSQATNDERNRQIA
ETGASLIKHW"
2986.3245
/notes="rnb transcription terminator"
BASE COUNT 793 a 798 c 869 g 807 t
ORIGIN

```

```

Query Match 100.0%; Score 1001; DB 12; Length 3267;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGGAGGATCATCAGCGGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 60
Db 1392 CGCTGGAGGATCATCAGCGGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 1333

Qy 61 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTTGGCGTCCGTTGGTTCGGTCAAT 120
Db 1332 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTTGGCGTCCGTTGGTTCGGTCAAT 1273

Qy 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCAGAGAGCGATAGAGCGGATGGCT 180
Db 1272 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCAGAGAGCGATAGAGCGGATGGCT 1213

Qy 181 GCGAATCGGAGCGGCGATACCGTAAGCAGAGAGCGGTACGCCCATTCGCGGCAA 240
Db 1212 GCGAATCGGAGCGGCGATACCGTAAGCAGAGAGCGGTACGCCCATTCGCGGCAA 1153

Qy 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCTGATAGCGGTCCGCCACACCA 300
Db 1152 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCTGATAGCGGTCCGCCACACCA 1093

Qy 301 GCGGCGCAGTGCATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 360
Db 1092 GCGGCGCAGTGCATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 1033

Qy 361 AGGCATCGCATGGTTCAGCAGAGATCTCCCGTTCGGGATGCGCGCTTGGACCTGG 420
Db 1032 AGGCATCGCATGGTTCAGCAGAGATCTCCCGTTCGGGATGCGCGCTTGGACCTGG 973

Qy 421 CGAACAGTTCGGTTCGGGAGCGCCCTGATGCTCTTCGTCAGATCATCTTCGATCGACAA 480
Db 972 CGAACAGTTCGGTTCGGGAGCGCCCTGATGCTCTTCGTCAGATCATCTTCGATCGACAA 913

Qy 481 GACCGGCTTCATCCAGATACGTCGCTCGATGCGATGTTTCGCTTGGTGGTTCGATG 540

```

```

Db 912 GACCGGCTTCCATCCGAGTAGCTGCTCGCTCGATGCGATGTTTCGTTGGTTCGATG 853
Qy 541 GCGAGGTAGCGGATCAAGCGTATGCGCGCGCGCATGTCATGATGATGATGATGATGAT 600
Db 852 GCGAGGTAGCGGATCAAGCGTATGCGCGCGCGCATGTCATGATGATGATGATGATGAT 793
Qy 601 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCTCTCCCGCGGACATTCGCCCAATAGCA 660
Db 792 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCTCTCCCGCGGACATTCGCCCAATAGCA 733
Qy 661 GCGAGTCCCTTCCCGCTTCAGTGAACAGTGCAGACAGTGCAGAGAGAGAGAGAGAGAGAG 720
Db 732 GCGAGTCCCTTCCCGCTTCAGTGAACAGTGCAGACAGTGCAGAGAGAGAGAGAGAGAG 673
Qy 721 TGGCCAGCACGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 780
Db 672 TGGCCAGCACGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 613
Qy 781 CGGTCTTGACAAAAGAACCGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 840
Db 612 CGGTCTTGACAAAAGAACCGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 553
Qy 841 AGCAGCCGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 552 AGCAGCCGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 493
Qy 901 GAGAACCTTGGTGCATTCATCTTGTTCATATGCGAAAGATGCTTCATCTTCCTTCCTTC 960
Db 492 GAGAACCTTGGTGCATTCATCTTGTTCATATGCGAAAGATGCTTCATCTTCCTTCCTTC 433
Qy 961 GATCAGATCTTGATCCCTCGCGCATCAGATCCTTCGCGCGC 1001
Db 432 GATCAGATCTTGATCCCTCGCGCATCAGATCCTTCGCGCGC 392

```

```

RESULT 8
A00782/c
LOCUS 3740 bp DNA linear PAT 16-FEB-1993
DEFINITION Plasmid pDm1.1 DNA sequence from patent EP0309746.
ACCESSION A00782
VERSION A00782.1 GI:14616
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 3740)
AUTHORS Certa,U.
TITLE Antimalaria vaccines
JOURNAL Patent: EP 0309746-A 11 05-APR-1989;
F. HOPMANN-LA ROCHE AG
FEATURES
location/Qualifiers
source
1. 3740
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN
Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGGAGGATCATCAGCGGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 60
Db 1290 CGCTGGAGGATCATCAGCGGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 1231

Qy 61 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTTGGCGTCCGTTGGTTCGGTTCGTCATT 120
Db 1230 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTTGGCGTCCGTTGGTTCGGTTCGTCATT 1171

Qy 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 180
Db 1170 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 1111

```



```

Db 390 GAGAACCTCGGTGCAATCCATCTTCTTCAATCATCGGAAACGATCCATCTCTCTCTT 331
QY 961 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 1001
Db 330 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 290

RESULT 10
A02297/c
LOCUS A02297 3740 bp DNA linear PAT 18-MAY-1993
DEFINITION Plasmid pDML1 DNA.
ACCESSION A02297
VERSION A02297.1 GI:345260
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences
REFERENCE 1 (bases 1 to 3740)
AUTHORS Certa, U., Gentz, R. and Takacs, B.
TITLE Plasmodium falciparum merozoite antigen peptides
JOURNAL Patent: EP 0283829-A 18 28-SEP-1988;
F. HOFFMANN-LA ROCHE AG
FEATURES
    Location/Qualifiers
    1..3740
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN
    Query Match 100.0%; Score 1001; DB 6; Length 3740;
    Best Local Similarity 100.0%; Pred. No. 0;
    Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGATCATCCAGCGCGTCCGGAACAGATTCGGAAGCCAACTTTTCATAG 60
Db 1290 CGCTGAGATCATCCAGCGCGTCCGGAACAGATTCGGAAGCCAACTTTTCATAG 1231
QY 61 AAGGGCGCGGTGGAATCGAATCTGTGATGCGAGGTTGGCGTCTGCTCGGTCAAT 120
Db 1230 AAGGGCGCGGTGGAATCGAATCTGTGATGCGAGGTTGGCGTCTGCTCGGTCAAT 1171
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGGATAGAGCGGATCGCT 180
Db 1170 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGGATAGAGCGGATCGCT 1111
QY 181 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 240
Db 1110 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 1051
QY 241 GCTCTTCAGCAATATCAGGGTAGCCAAACGCTATGTCTGATAGCGGTCCGCCACACCA 300
Db 1050 GCTCTTCAGCAATATCAGGGTAGCCAAACGCTATGTCTGATAGCGGTCCGCCACACCA 991
QY 301 GCGGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 360
Db 990 GCGGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 931
QY 361 AGGCATTCGCCATGCGGTACAGAGATCTCTCGCGTGGGATCGCGCTTTCAGCGCTG 420
Db 930 AGGCATTCGCCATGCGGTACAGAGATCTCTCGCGTGGGATCGCGCTTTCAGCGCTG 871
QY 421 CGAACAGTTCGGCTCGCGGAGCCCTCTGATGCTCTGTCAGATCATCTGATGACAA 480
Db 870 CGAACAGTTCGGCTCGCGGAGCCCTCTGATGCTCTGTCAGATCATCTGATGACAA 811
QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 540
Db 810 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 751
QY 541 GCGAGTACCGGATCAAGCGTATGTCAGCCCGCATTTGATCAGCCATGATGATCACTT 600

```

```

Db 750 GCGAGTAGCCGATCAAGCGTATGAGCCGCCCATTCATCAGCCATGATGATCTT 891
QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTCCGCCAATAGCA 660
Db 690 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTCCGCCAATAGCA 631
QY 661 GCCAGTCCCTTCCCGCTTTCAGTGCACAAAGTTCGAGCAGCTGCGGCAAGAACGCCCGTCG 720
Db 630 GCCAGTCCCTTCCCGCTTTCAGTGCACAAAGTTCGAGCAGCTGCGGCAAGAACGCCCGTCG 571
QY 721 TGGCCAGCAGATAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 570 TGGCCAGCAGATAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511
QY 781 CGGTCTTGACAAAAGAACCGCGCGCCCTGCGCTGACACAGCCCGGAAACACAGCGGCGCATCAG 840
Db 510 CGGTCTTGACAAAAGAACCGCGCGCCCTGCGCTGACACAGCCCGGAAACACAGCGGCGCATCAG 451
QY 841 AGCAGCGGATTTCTGTTGTCGCGAGTCATAGCGGATAGCTCTCTCAGCCAGCGCGCG 900
Db 450 AGCAGCGGATTTCTGTTGTCGCGAGTCATAGCGGATAGCTCTCTCAGCCAGCGCGCG 391
QY 901 GAGAACCTGCGTGCATCTCTTGTTCATCATGCGAAACGATCCTCATCTGCTCTCTT 960
Db 390 GAGAACCTGCGTGCATCTCTTGTTCATCATGCGAAACGATCCTCATCTGCTCTCTT 331
QY 961 GATCAGATCTTGCATCCCTGCGGATCAGATCTTGGCGGC 1001
Db 330 GATCAGATCTTGCATCCCTGCGGATCAGATCTTGGCGGC 290

RESULT 11
A07054/c
LOCUS A07054 3740 bp DNA linear PAT 25-AUG-1993
DEFINITION Artificial sequence neo and lacI gene for neomycin
phosphotransferase and lac repressor.
ACCESSION A07054
VERSION A07054.1 GI:412971
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3740)
AUTHORS Doebeli, H., Gentz, R. and Hochuli, E.
TITLE Homogeneous recombinant immune interferon fragments and
pharmaceutical compositions containing same
JOURNAL Patent: EP 0256424-A 11 24-FEB-1988;
F. HOFFMANN-LA ROCHE AG
FEATURES
    Location/Qualifiers
    1..3740
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN
    Query Match 100.0%; Score 1001; DB 6; Length 3740;
    Best Local Similarity 100.0%; Pred. No. 0;
    Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGTCCGGAACAGATTCGGAAGCCAACTTTTCATAG 60
Db 1290 CGCTGAGGATCATCCAGCGCGTCCGGAACAGATTCGGAAGCCAACTTTTCATAG 1231
QY 61 AAGGGCGCGGTGGAATCGAATCTGTGATGCGAGGTTGGCGTCTGCTCGGTCAAT 120
Db 1230 AAGGGCGCGGTGGAATCGAATCTGTGATGCGAGGTTGGCGTCTGCTCGGTCAAT 1171
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGGATAGAGCGGATCGCT 180
Db 1170 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGGATAGAGCGGATCGCT 1111
QY 181 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 240
Db 1110 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 1051
QY 241 GCTCTTCAGCAATATCAGGGTAGCCAAACGCTATGTCTGATAGCGGTCCGCCACACCA 300
Db 1050 GCTCTTCAGCAATATCAGGGTAGCCAAACGCTATGTCTGATAGCGGTCCGCCACACCA 991
QY 301 GCGGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 360
Db 990 GCGGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 931
QY 361 AGGCATTCGCCATGCGGTACAGAGATCTCTCGCGTGGGATCGCGCTTTCAGCGCTG 420
Db 930 AGGCATTCGCCATGCGGTACAGAGATCTCTCGCGTGGGATCGCGCTTTCAGCGCTG 871
QY 421 CGAACAGTTCGGCTCGCGGAGCCCTCTGATGCTCTGTCAGATCATCTGATGACAA 480
Db 870 CGAACAGTTCGGCTCGCGGAGCCCTCTGATGCTCTGTCAGATCATCTGATGACAA 811
QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 540
Db 810 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 751
QY 541 GCGAGTACCGGATCAAGCGTATGTCAGCCCGCATTTGATCAGCCATGATGATCACTT 600

```

Db 1110 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGAAAGCGGTGAGCCCATTCGCCGCA 1051
Qy 241 GCTCTTCAGCAATATACGGGTAGCCACGCTATGCTCTGATAGCGTCCGCCACACCCA 300
Db 1050 GCTCTTCAGCAATATACGGGTAGCCACGCTATGCTCTGATAGCGTCCGCCACACCCA 991
Qy 301 GCCGGCCACAGTCATGATATCAGAAAGCGGCATTTTCCACATGATATTCGGCAAGC 360
Db 990 GCCGGCCACAGTCATGATATCAGAAAGCGGCATTTTCCACATGATATTCGGCAAGC 931
Qy 361 AGGCATCGCCATGGGTACACACAGAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 420
Db 930 AGGCATCGCCATGGGTACACACAGAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 871
Qy 421 CGAACGTTTCGGTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTCTGATGACAA 480
Db 870 CGAACGTTTCGGTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTCTGATGACAA 811
Qy 481 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTCGTCAGATCATCTCTGATGACAA 840
Db 810 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTCGTCAGATCATCTCTGATGACAA 811
Qy 541 GGCAGTTCGGGATCAAGGATGAGAGCGCGCGATTCGATCAGCCATGATGATGAT 600
Db 750 GGCAGTTCGGGATCAAGGATGAGAGCGCGCGATTCGATCAGCCATGATGATGAT 691
Qy 601 TCTCGGAGCAGCAAGGTGAGATGACAGAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 660
Db 690 TCTCGGAGCAGCAAGGTGAGATGACAGAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 631
Qy 661 GCGAGTTCGGTTCGGTTCAGTACAAAGCGTTCGAGCAAGCGTTCGAGCAAGCGCGCTG 720
Db 630 GCGAGTTCGGTTCGGTTCAGTACAAAGCGTTCGAGCAAGCGTTCGAGCAAGCGCGCTG 571
Qy 721 TGSCCAGCAGTACGCGGTGCTCTCTGATGAGTTCATTCAGGCAACCGGACAGGT 780
Db 570 TGSCCAGCAGTACGCGGTGCTCTCTGATGAGTTCATTCAGGCAACCGGACAGGT 511
Qy 781 CGGTCTTGACAAAAGAACCGGCGCGCTGAGTACGAGCAAGCGTTCGAGCAAGCGCGCTG 840
Db 510 CGGTCTTGACAAAAGAACCGGCGCGCTGAGTACGAGCAAGCGTTCGAGCAAGCGCGCTG 451
Qy 841 AGCAGCGGATGCTCTGTTGCGGAGTACAGCGGATGCTCTCCACCCCAAGCGGCG 900
Db 450 AGCAGCGGATGCTCTGTTGCGGAGTACAGCGGATGCTCTCCACCCCAAGCGGCG 391
Qy 901 GAGAACTTCGTCGAATCCATCTTGTCAATCATCGGAAAGCATCTCTCATCTGCTCTT 960
Db 390 GAGAACTTCGTCGAATCCATCTTGTCAATCATCGGAAAGCATCTCTCATCTGCTCTT 331
Qy 961 GATCAGATCTTCATCCCTGCGCCATCAGATCTTTCGCGG 1001
Db 330 GATCAGATCTTCATCCCTGCGCCATCAGATCTTTCGCGG 290

RESULT 12

A10362/c

LOCUS

DEFINITION

Accession

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

A10362 Artificial sequence of plasmid pDM1.1. linear PAT 22-SEP-1993

A10362 Artificial sequence of plasmid pDM1.1. linear PAT 22-SEP-1993

A10362.1 GI:490684

synthetic construct

artificial sequences

1 (bases 1 to 3740)

Bujard, H. and Langer, M.

High repressible sequence for control of expression

Patent: EP 0303925-A 9 22-FEB-1989;

F. HOFFMANN-LA ROCHE AG

Location/Qualifiers

1..3740

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

BASE COUNT 813 a 1035 c 1056 g 836 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 1001; DB 6; Length 3740;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGAGGATCATCCAGCGCGCTCCGGAACGATTCGGAAGCCCAACCTTTTCATAG 60

Db 1290 CGCTGAGGATCATCCAGCGCGCTCCGGAACGATTCGGAAGCCCAACCTTTTCATAG 1231

Qy 61 AAGCGCGGTGGAAATCGAATCTCGGTATGAGGAGGTTGGCGCTTGGTTCGGTCAAT 120

Db 1230 AAGCGCGGTGGAAATCGAATCTCGGTATGAGGAGGTTGGCGCTTGGTTCGGTCAAT 1171

Qy 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATGAGAGCGGATGCGCT 180

Db 1170 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATGAGAGCGGATGCGCT 1111

Qy 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCA 240

Db 1110 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCA 1051

Qy 241 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCGTATGCTCTGATAGCGGTCCGCAACCCA 300

Db 1050 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCGTATGCTCTGATAGCGGTCCGCAACCCA 991

Qy 301 GCGGCGCACAGTGCATGATTCAGAAAGCGGCAATTTCCACCATGATATTCGCAAGC 360

Db 990 GCGGCGCACAGTGCATGATTCAGAAAGCGGCAATTTCCACCATGATATTCGCAAGC 931

Qy 361 AGGCATCGCCATGGGTACAGACAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 420

Db 930 AGGCATCGCCATGGGTACAGACAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 871

Qy 421 CGAACGTTTCGGTGGCGGAGCGCTGATGCTCTTCGTCAGATCATCTCTGATGACAA 480

Db 870 CGAACGTTTCGGTGGCGGAGCGCTGATGCTCTTCGTCAGATCATCTCTGATGACAA 811

Qy 481 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTCGTCAGATCATCTCTGATGACAA 540

Db 810 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTCGTCAGATCATCTCTGATGACAA 751

Qy 541 GGCAGTTCGGGATCAAGGATGAGAGCGCGGATTCGATCAGCCATGATGATGAT 600

Db 750 GGCAGTTCGGGATCAAGGATGAGAGCGCGGATTCGATCAGCCATGATGATGAT 691

Qy 601 TCTCGGAGCAGCAAGGTGAGATGACAGAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 660

Db 690 TCTCGGAGCAGCAAGGTGAGATGACAGAGATCTTCGCCGTGCGGCATGCGGCTTTCAGCGT 631

Qy 661 GCGAGTTCGGTTCGGTTCAGTACAAAGCGTTCGAGCAAGCGTTCGAGCAAGCGCGCTG 720

Db 630 GCGAGTTCGGTTCGGTTCAGTACAAAGCGTTCGAGCAAGCGTTCGAGCAAGCGCGCTG 571

Qy 721 TGSCCAGCAGTACGCGGTGCTCTCTGATGAGTTCATTCAGGCAACCGGACAGGT 780

Db 570 TGSCCAGCAGTACGCGGTGCTCTCTGATGAGTTCATTCAGGCAACCGGACAGGT 511

Qy 781 CGGTCTTGACAAAAGAACCGGCGCGCTGAGTACGAGCAAGCGTTCGAGCAAGCGCGCTG 840

Db 510 CGGTCTTGACAAAAGAACCGGCGCGCTGAGTACGAGCAAGCGTTCGAGCAAGCGCGCTG 451

Qy 841 AGCAGCGGATGCTCTGTTGCGGAGTACAGCGGATGCTCTCCACCCCAAGCGGCG 900

Db 450 AGCAGCGGATGCTCTGTTGCGGAGTACAGCGGATGCTCTCCACCCCAAGCGGCG 391

Qy 901 GAGAACTTCGTCGAATCCATCTTGTCAATCATCGGAAAGCATCTCTCATCTGCTCTT 960

Db 390 GAGAACTTCGTCGAATCCATCTTGTCAATCATCGGAAAGCATCTCTCATCTGCTCTT 331

```
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 1001
Db 330 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 290

RESULT 13
AL14595/c
LOCUS Al14595 Synthetic nucleotide sequence of the pD1,1 plasmid. circular PAT 21-FEB-1994
DEFINITION Al14595
ACCESSION Al14595
VERSION Al14595.1 GI:491836
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 3740)
AUTHORS Bannwarth, W., Certa, U., Mous, J. and Stueber, D.
TITLE Polypeptides that elicit antibodies against AIDS virus
JOURNAL Patent: EP 0219106-A 8 22-APR-1987;
F. HOFFMANN-LA ROCHE AG
FEATURES
    source
        Location/Qualifiers
            1..3740
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
                357..1151
                    /codon_start=1
                    /transl_table=1
                    /product="neomycin phosphotransferase"
                    /protein_id="CAA01174.1"
                    /db_xref="GI:491837"
                    /translation="MIEQDLHAGSPAAWVRELFQVDWAQQTIGSDAAVFLSAOGR
                    PVLVKTDLSGALNELDEARLSHLATGVPCAAVLVVTAEGRDMLLGEVFGQDL
                    LSHLAPAEKVSIMADAMRLRLHLPATCFPDQAKHRIERARTMERAGLVQDDLDE
                    EQGLAPAEFLARLAKRMPDGLVTHGDACLPMVNGRFGSFGIDCGRLGVADRY
                    OIALATRDIAELGGEWADFLVLYGIAAPDSORIAFYRLIDFF"
                    /codon_start=1
                    /transl_table=1
                    /product="lac repressor"
                    /protein_id="CAA01175.1"
                    /db_xref="GI:4526670"
                    /translation="MKPVTLYDAEYAGVSYQTVSRVVNOASHVSAKTREKVEAAMAE
                    LNYINRVAQQLAGKSLILIGVATSLALHAPSQIVAAIKSRADQLGASVVSVMS
                    GVEACKAAVHNLAAQVSLIINVLDDQDAIAVEAACTNPALPLDVSQDTPINSII
                    FSHEDGTRIGVHLYALGHQIALLAGPLSSVSLRLAGHKLTRNQIPIAREG
                    DRMSGFGQTMQNLNEGIVPTAMLVANDQVALGAMRAITSEGLRVGADISVVGVDYD
                    EDSSCTIPPLTTIKQDFRLGQTSVDRLLQLSQQAQVKGQNLPLVSLVKKRTTLAPNT
                    QTSAPRALADSLMQLARQVSRLESQ"
BASE COUNT 813 a 1036 c 1055 g 836 t
ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCCGAAGAACGATCCGAGCCCAACCTTTCATAG 60
Db 1290 CGCTGAGGATCATCCAGCGCGCTCCCGAAGAACGATCCGAGCCCAACCTTTCATAG 1231

QY 61 AAGCGCGCGTGGATCGAAATCTCGTATGACGAGTTGGCGTGCCTGTGTCGTCATT 120
Db 1230 AAGCGCGCGTGGATCGAAATCTCGTATGACGAGTTGGCGTGCCTGTGTCGTCATT 1171

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCTGTCAGAGGCGGATAGAGCGGATGCGCT 180
Db 1170 TCGAACCCAGAGTCCCGCTCAGAGAACTCTGTCAGAGGCGGATAGAGCGGATGCGCT 1111

QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGACGCCCATTCGCGGCCAA 240
Db 1110 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGACGCCCATTCGCGGCCAA 1051

RESULT 14
AL19608/c
LOCUS Al19608 Artificial sequence of plasmid pD1,1. linear PAT 09-MAY-1994
DEFINITION Al19608
ACCESSION Al19608
VERSION Al19608.1 GI:513600
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 3740)
AUTHORS Bannwarth, W., Caspers, P., Le Grice, S. and Mous, J.
TITLE Recombinant HIV-2 polypeptides
JOURNAL Patent: EP 0316695-A 6 24-MAY-1989;
F. HOFFMANN-LA ROCHE AG
FEATURES
    source
        Location/Qualifiers
            1..3740
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"

QY 241 GCTCTTACGCAATATCACGGGTAGCCCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 300
Db 1050 GCTCTTACGCAATATCACGGGTAGCCCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 991

QY 301 GCCGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 990 GCCGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 931

QY 361 AGGCATCGCCATGGGTTCACGACGAGATCTTCGCCGTGGGCATCGCGCCTTGAGCCTGG 420
Db 930 AGGCATCGCCATGGGTTCACGACGAGATCTTCGCCGTGGGCATCGCGCCTTGAGCCTGG 871

QY 421 CGAACAGTTGGCTGGCGCGAGCCCTTGATGCTCTTCTCCAGATCATCTGATGACGACAA 480
Db 870 CGAACAGTTGGCTGGCGCGAGCCCTTGATGCTCTTCTCCAGATCATCTGATGACGACAA 811

QY 481 GACCGGCTTCATCCGAGTACGCTCGCTCGATCGATGCTTTCGCTTGGTGGTGGATG 540
Db 810 GACCGGCTTCATCCGAGTACGCTCGCTCGATCGATGCTTTCGCTTGGTGGTGGATG 751

QY 541 GGCAGGTAGCCGGATCAAGCGTATGCAAGCGCGCCGATTTGCATCAGCCATGATGATCTT 600
Db 750 GGCAGGTAGCCGGATCAAGCGTATGCAAGCGCGCCGATTTGCATCAGCCATGATGATCTT 691

QY 601 TCTCGGACGAGCAGGTGAGTACAGAGAGATCTTCCCGCGGACATTCGCCCCCAATAGCA 660
Db 690 TCTCGGACGAGCAGGTGAGTACAGAGAGATCTTCCCGCGGACATTCGCCCCCAATAGCA 631

QY 661 GCCAGTCCCTTCCCGCTTCAAGCGTATGCAAGCGCGCCGATTTGCATCAGCCATGATGATCTT 720
Db 630 GCCAGTCCCTTCCCGCTTCAAGCGTATGCAAGCGCGCCGATTTGCATCAGCCATGATGATCTT 571

QY 721 TGGCCACGACGATAGCCGCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 570 TGGCCACGACGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511

QY 781 CGGTCTTTGACAAAAGAACCGCGCGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 510 CGGTCTTTGACAAAAGAACCGCGCGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451

QY 841 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 450 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391

QY 901 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 390 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 331

QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 1001
Db 330 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 290
```



```
Qy 661 GCCAGTCCCTTCCCGCTTTCAGTGACAAACGTCCAGCACAGCTGCGCAAGGAACGCCCGTCG 720
Db |||||
Qy 630 GCCAGTCCCTTCCCGCTTTCAGTGACAAACGTCCAGCACAGCTGCGCAAGGAACGCCCGTCG 571
Db |||||
Qy 721 TGGCCAGCCACGATAGCCGCGCTGCTCGTCTCTGCAAGTTTCATTCAGGGGACCCGGACAGGT 780
Db |||||
Qy 570 TGGCCAGCCACGATAGCCGCGCTGCTCGTCTCTGCAAGTTTCATTCAGGGGACCCGGACAGGT 511
Db |||||
Qy 781 CGGTCTTTGACAAAAGAACCGGCGGCCCTCGCTGACAGCGGGAACACGGCGGCATCAG 840
Db |||||
Qy 510 CGGTCTTTGACAAAAGAACCGGCGGCCCTCGCTGACAGCGGGAACACGGCGGCATCAG 451
Db |||||
Qy 841 AGCAGCCGATTTGCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAGCGGCG 900
Db |||||
Qy 450 AGCAGCCGATTTGCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAGCGGCG 391
Db |||||
Qy 901 GAGAACTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCCTGTCTCTT 960
Db |||||
Qy 390 GAGAACTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCCTGTCTCTT 331
Db |||||
Qy 961 GATCAGATTTGATCCCTCGGCCATCAGATCCTTGGCGGC 1001
Db |||||
Qy 330 GATCAGATTTGATCCCTCGGCCATCAGATCCTTGGCGGC 290
Db |||||
```

Search completed: February 16, 2004, 01:25:58
Job time : 3968.99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:07:11 ; Search time 324.065 Seconds
(without alignments)
8338.264 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 cgcggaggatccagcc.....gccatcagatccttgggcgc 1001

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

```
1: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-embl/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-embl/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	1001	100.0	2220	21	AAZ58064	
C	2	1001	100.0	3176	14	AAQ42402
C	3	1001	100.0	3176	20	AAQ01440
C	4	1001	100.0	3177	13	AAQ27504
C	5	1001	100.0	3740	9	AA080959
C	6	1001	100.0	3740	9	AA081154
C	7	1001	100.0	3740	9	AA080483
C	8	1001	100.0	3740	11	AAQ06307
						Vector plasmid pUN
						Tn5seq1 transposon
						Tn5seq1 transposon
						Transposon Tn5seq1
						Plasmid pDM1.1 enc
						Plasmid pDM1.1. P
						pDM1.1 contg. of the
						Sequence of plasmid

Vector plasmid pUN
Tn5seq1 transposon
Tn5seq1 transposon
Transposon Tn5seq1
Plasmid pDML1 enc
Plasmid pDML1, P
pDML1 contg. the
Sequence of plasmid

C	9	1001	100.0	3740	12	AAQ12786
C	10	1001	100.0	3740	14	AAQ34610
	11	1001	100.0	3974	19	AAV33721
	12	1001	100.0	3974	19	AAV66372
	13	1001	100.0	3974	19	AAV56699
	14	1001	100.0	3974	19	AAV57971
	15	1001	100.0	3974	19	AAV42677
	16	1001	100.0	3974	20	AAZ32002
	17	1001	100.0	3974	20	AAZ19915
	18	1001	100.0	3974	20	AAZ10532
	19	1001	100.0	3974	20	AAZ11171
	20	1001	100.0	3974	20	AAZ87252
	21	1001	100.0	3974	20	AAZ13894
	22	1001	100.0	3974	20	AAZ08956
	23	1001	100.0	3974	21	AAZ75540
	24	1001	100.0	3974	21	AAZ72757
	25	1001	100.0	3974	21	AAZ48490
	26	1001	100.0	3974	21	AAZ71310
	27	1001	100.0	3974	21	AAZ27353
	28	1001	100.0	3974	22	AAZ32008
	29	1001	100.0	3974	22	AAZ54432
	30	1001	100.0	3974	24	AAZ54432
	31	1001	100.0	3974	24	ABQ73147
	32	1001	100.0	3974	24	ABQ10061
	33	1001	100.0	3974	24	ABQ23719
	34	1001	100.0	3974	25	ABQ17117
	35	1001	100.0	3974	25	ABQ76975
	36	1001	100.0	3974	25	AAZ49531
	37	1001	100.0	3974	25	ABQ57778
	38	1001	100.0	3974	25	ABQ83101
	39	1001	100.0	3984	20	AAZ33990
	40	1001	100.0	3997	24	ABQ78682
C	41	1001	100.0	4208	19	AAZ4537
	42	1001	100.0	4208	20	AAZ33993
	43	1001	100.0	4208	22	AAZ82961
	44	1001	100.0	4256	19	AAZ63305
	45	1001	100.0	4256	22	AAZ21539

ALIGNMENTS

RESULT 1
AAZ58064
ID AAZ58064 standard; DNA; 2220 BP.

XX AC AAZ58064;

XX AC AC

DT 25-APR-2000 (first entry)

XX Vector plasmid pUNI-10.

DE Plasmid pUNI-10; vector; site-specific recombination; subcloning;

DE Univector Fusion System; Univector Plasmid-Fusion System; ss.

KW Chimeric - Escherichia coli.

XX Chimeric - Bos taurus.

OS Chimeric - Bacteriophage T7

XX Chimeric - Synthetic.

XX Key Location/Qualifiers

FT rep_origin 1..400

FT /*tag= a

FT /*note= "R6K-gamma origin of replication"

FT misc_feature 401..414

FT /*tag= b

FT misc_feature 415..448

FT /*tag= c

FT /*note= "polylinker facilitating lox site exchange"

FT misc_feature 449..527

FT /*tag= d

FT /*note= "wild-type loxP recombinase target site"

FT /*tag= d

FT /*note= "polylinker used as DNA insertion site"

FT

FT polyA_site 528..750
 FT /*tag= e
 FT /note= "BGH polyA site"
 FT terminator 751..890
 FT /*tag= f
 FT /note= "T7 terminator"
 FT prim_transcript 896..2220
 FT /*tag= g
 FT /note= "kanamycin resistance gene from Tn5"
 PN WO200005355-A1.
 PN 03-FEB-2000.
 PD 27-JUL-1998; 98WO-US15481.
 PF 24-JUL-1998; 98US-0122384.
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 XX Elledge SJ, Liu Q;
 XX WPI; 2000-171429/15.
 XX New methods for recombination of nucleic acid constructs, useful for
 PT the rapid subcloning of nucleic acid sequences in vivo and in vitro
 XX Claim 44; Page 63-64; 110pp; English.
 CC The present sequence is that of novel vector plasmid pUNI-10. It
 CC includes: the conditional origin of replication from RepK-gamma,
 CC derived from plasmid RepK (ATCC 37120); a NotI-XpnI polylinker that
 CC facilitates the exchange of lox sites; a wild-type loxP site (see
 CC also AA258066); a polylinker used for the insertion of genomic DNA
 CC or cDNA sequences; a bovine growth hormone polyA site; a T7
 CC terminator; and a kanamycin resistance selectable marker gene from
 CC transposon Tn5, modified to remove the native NcoI site. pUNI-10
 CC is an example of novel Univectors or pUNI vectors of the invention
 CC that have a sequence-specific recombinase target site (e.g. loxP)
 CC preceding the insertion site for the gene of interest, a selectable
 CC marker gene (optional) and a conditional origin of replication that
 CC is active only in host cells expressing the requisite transacting
 CC replication factor (optional). The vectors are designed to contain
 CC a gene of interest but to lack a promoter for expression of the
 CC gene. The vectors are used in a novel method for the rapid
 CC subcloning of nucleic acid sequences in vivo and in vitro without
 CC the need of restriction endonucleases. The method is referred to
 CC as the Univector Fusion System or Univector Plasmid-fusion system
 CC (UPS). The UPS uses site-specific recombination to catalyze plasmid
 CC fusion between a Univector and host vectors containing regulatory
 CC information. In some embodiments, plasmid fusion events are
 CC genetically selected and result in placement of the gene of interest
 CC under the control of novel regulatory elements. A second UPS-related
 CC method allows for the precise transfer of coding sequences alone from
 CC a Univector into a host vector. UPS further provides means for the
 CC subcloning of entire nucleic acid libraries and the directional
 CC cloning of linear nucleic acid molecules, e.g. PCR products.
 XX
 XX Sequence 2220 BP; 506 A; 610 C; 581 G; 523 T; 0 other;
 Query Match 100.0%; Score 1001; DB 21; Length 2220;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAAGAGATTCCGAAGCCCAACCTTTCATAG 60
 Db 932 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAAGAGATTCCGAAGCCCAACCTTTCATAG 991
 QY 61 AAGCGCGCGGTGGAATCGAAATCTCGTATGCGAGGTGGCGGTGCTGTGGTGCATTT 120
 Db 992 AAGCGCGCGGTGGAATCGAAATCTCGTATGCGAGGTGGCGGTGCTGTGGTGCATTT 1051
 QY 121 TCGAACCACCGAGTCCCGGTCTCAGAGAACTCGTCAAGAGGGGATAGAGGCGATGCGCT 180

Db 1052 TCGAACCACCGAGTCCCGGTCTCAGAGAACTCGTCAAGAGCGGATAGAGCGCGATGCGCT 1111
 QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAAAGCGGTGAGCCCATTCGCGGCA 240
 Db 1112 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAAAGCGGTGAGCCCATTCGCGGCA 1171
 QY 241 GCTCTTCAGCAATATACAGCGGTAGCCAAACGCTATGCTCTGATACGCGTCCGCGCACACCA 300
 Db 1172 GCTCTTCAGCAATATACAGCGGTAGCCAAACGCTATGCTCTGATACGCGTCCGCGCACACCA 1231
 QY 301 GCGGCGCACAGTGTGATGATTCAGAAAGAGCGGCAATTTCCACATGATATTCGGAAGC 360
 Db 1232 GCGGCGCACAGTGTGATGATTCAGAAAGAGCGGCAATTTCCACATGATATTCGGAAGC 1291
 QY 361 AGGCATCGCATGGTCAACGAGAGATCTCTCGCGTCCGCGCATGCGCGCTTCAGGCTGG 420
 Db 1292 AGGCATCGCATGGTCAACGAGATCTCTCGCGTCCGCGCATGCGCGCTTCAGGCTGG 1351
 QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCCAGATCATCTGATGACAA 480
 Db 1352 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCCAGATCATCTGATGACAA 1411
 QY 481 GACCGGCTTCATCCCGAGTACGCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 540
 Db 1412 GACCGGCTTCATCCCGAGTACGCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 1471
 QY 541 GGCAGGTAGCCGGATCAAGCGTATGACGCGCGCCGCAATTCGATCAGCGCATGATCGATAC 600
 Db 1472 GGCAGGTAGCCGGATCAAGCGTATGACGCGCGCCGCAATTCGATCAGCGCATGATGATAC 1531
 QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTCTGCGCGGCACTTCGCCCAATAGCA 660
 Db 1532 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTCTGCGCGGCACTTCGCCCAATAGCA 1591
 QY 661 GCCAGTCCCTTCGCGCTTCAGTCAACAGTTCGAGCACAGCTGGCGCAAGAAACGCCGCTCG 720
 Db 1592 GCCAGTCCCTTCGCGCTTCAGTCAACAGTTCGAGCACAGCTGGCGCAAGAAACGCCGCTCG 1651
 QY 721 TGGCGAGCGACGATAGCGCGCTCGCTCGTCTGCAAGTTCATTCAGGGGACCGGACAGGT 780
 Db 1652 TGGCGAGCGACGATAGCGCGCTCGCTCGTCTGCAAGTTCATTCAGGGGACCGGACAGGT 1711
 QY 781 CGGTCTTGACAAAAGAACCGCGCGCCCTCGCTCGACAGCGGCAACACGCGGCGCATCAG 840
 Db 1712 CGGTCTTGACAAAAGAACCGCGCGCCCTCGCTCGACAGCGGCAACACGCGGCGCATCAG 1771
 QY 841 AGCAGCGGATGTCTGTGTGCGCCAGTCTATAGCCGAATAGCCTCTCCACCCAAAGCGCGCG 900
 Db 1772 AGCAGCGGATGTCTGTGTGCGCCAGTCTATAGCCGAATAGCCTCTCCACCCAAAGCGCGCG 1831
 QY 901 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAAGCATCTCATCTCTGTCTTT 960
 Db 1832 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAAGCATCTCATCTCTGTCTTT 1891
 QY 961 GATCAGATCTTGATCCCTTCGCGCATCAGATCTTCCTGGCGGC 1001
 Db 1892 GATCAGATCTTGATCCCTTCGCGCATCAGATCTTCCTGGCGGC 1932

RESULT 2
 AAQ42402/c
 ID AAQ42402 standard; DNA; 3176 BP.
 XX
 AC AAQ42402;
 XX
 DT 25-MAR-2003 (updated)
 DT 13-SEP-1993 (first entry)
 XX
 DE Tn5seq1 transposon sequence.
 XX
 KW Bidirectional sequencing; restriction enzyme cleavage sites; marker;
 Tn5 transposase gene; E. coli; ss.

```

XX OS Synthetic.
XX FT
XX FT
XX FT
XX FT Key
XX FT misc_RNA
XX FT
XX FT Location/Qualifiers
XX FT 1..19
XX FT /tag= a
XX FT /note= "O-end Tn5 transposon sequence"
XX FT complement (1..29)
XX FT /tag= b
XX FT /note= "T7 RNA"
XX FT 21..26
XX FT /tag= c
XX FT /note= "BamHI restriction site"
XX FT 376..381
XX FT /tag= d
XX FT /note= "BclI restriction site"
XX FT 407..420
XX FT /tag= e
XX FT /note= "codes for kanamycin selectable marker"
XX FT 932..937
XX FT /tag= f
XX FT /note= "BssHI restriction site"
XX FT 935..941
XX FT /tag= g
XX FT /note= "SphI restriction site"
XX FT 967..972
XX FT /tag= h
XX FT /note= "NcoI restriction site"
XX FT 1051..1055
XX FT /tag= i
XX FT /note= "AvaII restriction site"
XX FT 1216..1222
XX FT /tag= j
XX FT /note= "AseI restriction site"
XX FT 1372..1377
XX FT /tag= k
XX FT /note= "SmaI restriction site"
XX FT 1388..1391
XX FT /tag= l
XX FT /note= "NruI restriction site"
XX FT 1540..1545
XX FT /tag= m
XX FT /note= "SalI restriction site"
XX FT complement (1549..2979)
XX FT /tag= n
XX FT /note= "tnp (Tn5 transposase gene)"
XX FT 1871..1876
XX FT /tag= o
XX FT /note= "HindIII restriction site"
XX FT 2051..2056
XX FT /tag= p
XX FT /note= "NheI restriction site"
XX FT 2256..2261
XX FT /tag= q
XX FT /note= "BamI restriction site"
XX FT 2361..2366
XX FT /tag= r
XX FT /note= "Eco4VII restriction site"
XX FT 2471..2478
XX FT /tag= s
XX FT /note= "NotI restriction site"
XX FT 2581..2586
XX FT /tag= t
XX FT /note= "XhoI restriction site"
XX FT 2611..2617
XX FT /tag= u
XX FT /note= "SacII restriction site"
XX FT 2856..2861
XX FT /tag= v
XX FT /note= "SspI restriction site"
XX FT 2881..2886
XX FT /tag= w
XX FT /note= "HpaI restriction site"
XX FT

FT misc_RNA 3042..3157
FT /tag= x
FT /note= "SP6 RNA"
FT 3142..3147
FT /tag= y
FT /note= "BamHI site"
FT 3158..3176
FT /tag= z
FT /note= "O-end of Tn5 transposon"
XX FT
XX FT US5212080-A.
XX FT
XX FT 18-MAY-1993.
XX FT
XX FT 12-APR-1990; 90US-0508379.
XX FT
XX FT 05-OCT-1987; 87US-0105422.
XX FT 12-APR-1990; 90US-0508379.
XX FT
XX FT (UNIW ) UNIV WASHINGTON.
XX FT
XX FT Berg DE, Huang HV, Nag DK;
XX FT MPI; 1993-175455/21.
XX FT
XX FT Bidirectional sequencing of long DNA molecules - by inserting
XX FT transposon Tn5seq1 into DNA molecule of E coli cell
XX FT
XX FT Disclosure; Fig 4; 17pp; English.
XX FT
XX FT The sequence shown is that of a novel transposon useful for
XX FT sequencing long DNA sequences. The transposon comprises a partial
XX FT sequence of transposon Tn5 with the oligonucleotide primers from
XX FT phages SP6 and T7 inserted near the opposite ends of Tn5. At each
XX FT terminus a segment of 19 nucleotides of either the O-end or I-end
XX FT of Tn5 are present and restriction enzyme sites are positioned
XX FT throughout the sequence (esp. with one at either end, less than
XX FT 20 bases from each terminal segment. Segments of foreign DNA of
XX FT at least 17 bases are positioned directly adjacent to each restriction
XX FT site. The transposon also comprises a Tn5 transposase gene insert and
XX FT a kanamycin selectable marker for the transposable DNA unit.
XX FT (Updated on 25-MAR-2003 to correct PF field.)
XX FT
XX FT Sequence 3176 BP; 661 A; 898 C; 893 G; 724 T; 0 other;
XX FT
XX FT Query Match 100.0%; Score 1001; DB 14; Length 3176;
XX FT Best Local Similarity 100.0%; Pred. No. 0;
XX FT Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX FT
XX FT 1 CGCTGGAGGATCATCCAGCGCGTCCCGGAAACGATTCGAGCCCAACCTTTCATAG 60
XX FT 1340 CGCTGGAGGATCATCCAGCGCGTCCCGGAAACGATTCGAGCCCAACCTTTCATAG 1281
XX FT
XX FT 61 AAGGCGGCGTGGATCGAATCTCGTATCGCAGGTTGGGCGTGGTGGTGGTCAATT 120
XX FT 1280 AAGGCGGCGTGGATCGAATCTCGTATCGCAGGTTGGGCGTGGTGGTGGTCAATT 1221
XX FT
XX FT 121 TCGAACCACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAGCGGATCGGCT 180
XX FT 1220 TCGAACCACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAGCGGATCGGCT 1161
XX FT
XX FT 181 GCGAATCGGAGCGCGGATACCGTAAAGCAAGAGAGCGGTCAGCCCATTCGCGGCAA 240
XX FT 1160 GCGAATCGGAGCGCGGATACCGTAAAGCAAGAGAGCGGTCAGCCCATTCGCGGCAA 1101
XX FT
XX FT 241 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCA 300
XX FT 1100 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCA 1041
XX FT
XX FT 301 GCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
XX FT 1040 GCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 981

```

QY 361 AGGATCCGATCGGTCACACGAGATCCTCGCGTCGGGATGCGGCTTGAGCCTGG 420
 Db 980 AGGATCCGATCGGTCACACGAGATCCTCGCGTCGGGATGCGGCTTGAGCCTGG 921
 QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTTCCTCCAGATCATCTGATCGACAA 480
 Db 920 CGAACAGTTCGGTGGCGGAGCCCTGATGCTTCCTCCAGATCATCTGATCGACAA 861
 QY 481 GACCGGTTCCATCCGAGTACGTGCTGCTCGATGCGATGTTTCGGTTCGATG 540
 Db 860 GACCGGTTCCATCCGAGTACGTGCTGCTCGATGCGATGTTTCGGTTCGATG 801
 QY 541 GCGAGTACCGGATCAAGGATGATGAGCGCGCATTCATCAGCCATGATGATCTT 600
 Db 800 GCGAGTACCGGATCAAGGATGATGAGCGCGCATTCATCAGCCATGATGATCTT 741
 QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTTGCCTCGGCACTTCGCCAATAGCA 660
 Db 740 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTTGCCTCGGCACTTCGCCAATAGCA 681
 QY 661 GCCAGTCCCTCCGCTTCAGTGACACAGTGCAGCAGCTGGCGAAGAAAGCCGCTG 720
 Db 680 GCCAGTCCCTCCGCTTCAGTGACACAGTGCAGCAGCTGGCGAAGAAAGCCGCTG 621
 QY 721 TGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 620 TGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
 QY 781 CGGCTTGACAAAAGAACCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 560 CGGCTTGACAAAAGAACCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
 QY 841 AGCAGCGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 500 AGCAGCGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 QY 901 GAGAACTGCTGATTCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
 Db 440 GAGAACTGCTGATTCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 381
 QY 961 GATCAGATCTGATTCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 Db 380 GATCAGATCTGATTCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340

RESULT 3

AAX01440/c
 ID AAX01440 standard; DNA; 3176 BP.

AC AAX01440;

XX 28-APR-1999 (first entry)

XX Tn5seq1 transposon.

XX Tn5seq1 transposon; RNA transcription; gene hyperexpression;
 KW strong promoter; Sp6 promoter; T7 promoter; ss.

XX Synthetic.

XX Key Location/Qualifiers

XX misc_feature 1500..1500

FT /*tag= a

FT /note= "specifically claimed region (claim 1)"

FT /*tag= b

FT /note= "O-end"

FT 3158..3176

FT /*tag= c

FT /note= "O-end"

XX US5869296-A.

XX

PD 09-FEB-1999.
 PF 14-JAN-1993; 93US-0004406.
 XX
 PR 05-OCT-1987; 87US-0105422.
 PR 12-APR-1990; 90US-0508382.
 PR 14-JAN-1993; 93US-0004406.
 XX (UNIW) UNIV WASHINGTON.
 PA
 PI Beig DE, Huang HV, Nag DK;
 XX WPI; 1999-152772/13.
 DR Obtaining hyperexpression of genes in Escherichia coli hosts - by
 XX insertion of transposon Tn5seq1 such that the strong Sp6 and T7
 PT promoters are adjacent to the host genes
 PS Claim 1; Fig 4a-d; 19pp; English.
 XX This sequence represents the Tn5seq1 transposon used in the method of the
 CC invention. The method is for RNA transcription, and comprises the
 CC insertion of the Tn5seq1 transposon into an E. coli DNA molecule to
 CC obtain hyperexpression of genes adjacent to strong promoters Sp6 or T7.
 CC The transposon is useful for stimulating the transcription of genes
 CC adjacent to the heterologous Sp6 or T7 promoters in E. coli, for making
 CC RNA transcripts in vitro and the hyperexpression of specific
 CC transcription of genes (adjacent to the Sp6 or T7 ends) in vivo. Tn5seq1
 CC offers a less laborious method of sequencing long DNA molecules than
 CC current methods such as base-specific chemical cleavage and enzymatic
 CC chain termination.
 XX
 SQ Sequence 3176 BP; 661 A; 898 C; 893 G; 724 T; 0 other;
 Query Match 100.0%; Score 1001; DB 20; Length 3176;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60
 Db 1340 CGCTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 1281
 QY 61 AAGCGCGGTGGAATCGAAATCTCGTATGGCAGGTGGCGGTGGTGGTGGTGGTGGT 120
 Db 1280 AAGCGCGGTGGAATCGAAATCTCGTATGGCAGGTGGCGGTGGTGGTGGTGGTGGT 1221
 QY 121 TCGAACCCCGAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGATGCGCT 180
 Db 1220 TCGAACCCCGAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGATGCGCT 1161
 QY 181 GCGAATCGGAGCGGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCAA 240
 Db 1160 GCGAATCGGAGCGGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCAA 1101
 QY 241 GCTCTTCAGCAATATCAGCGGTAGCCACGCTATGCTGATAGCGGTCCGCCACACCCA 300
 Db 1100 GCTCTTCAGCAATATCAGCGGTAGCCACGCTATGCTGATAGCGGTCCGCCACACCCA 1041
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGCAAGC 360
 Db 1040 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGCAAGC 981
 QY 361 AGGATCCCATGCGGTACGACGAGATCTTCGCCGTGCGGCGATGCGGCTTGAGCCTGG 420
 Db 980 AGGATCCCATGCGGTACGACGAGATCTTCGCCGTGCGGCGATGCGGCTTGAGCCTGG 921
 QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTTCCTCGTCCAGATCATCTGATCGACAA 480
 Db 920 CGAACAGTTCGGTGGCGGAGCCCTGATGCTTCCTCGTCCAGATCATCTGATCGACAA 861
 QY 481 GACCGGTTCCATCCGAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 860 GACCGGTTCCATCCGAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801

```
QY 541 GGCAGTAGCCGAGTCAAGCGTATGACGCGCGCGCATTCGATCAGCCATGATGGATCTT 600
Db 800 GGCAGTAGCCGAGTCAAGCGTATGACGCGCGCGCATTCGATCAGCCATGATGGATCTT 741
QY 601 TCTCGGAGGAGCAAGGTGAGATGACGAGAGATCCTGCGCGCGCATTCGCGCCATAGCA 660
Db 740 TCTCGGAGGAGCAAGGTGAGATGACGAGAGATCCTGCGCGCGCATTCGCGCCATAGCA 681
QY 661 GCGAGTCCCTTCCGCTTCAGTGAACAGTGCAGACAGCTGCGAAGAAAGCGCCGTCG 720
Db 680 GCGAGTCCCTTCCGCTTCAGTGAACAGTGCAGACAGCTGCGAAGAAAGCGCCGTCG 621
QY 721 TGSCCAGCCAGATAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 620 TGSCCAGCCAGATAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
QY 781 CGTCTTGTGAAAAAGAACCGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 560 CGTCTTGTGAAAAAGAACCGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
QY 841 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 500 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
QY 901 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 440 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
QY 961 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 1001
Db 380 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 340
```

RESULT 4

AAQ27504/c
ID AAQ27504 standard; DNA; 3177 BP.

XX AAQ27504;

XX AAQ27504;

DT 10-MAR-2003 (updated)

DT 05-FEB-1993 (first entry)

DE Transposon Tn5seq1.

XX Transposon Tn5seq1.

XX Transposon Tn5seq1.

KW Phase T7; RNA polymerase; DNA sequencing; ss.

OS Transposition; O-end sequence; insertional hotspot; Phage SP6;

OS Phase T7; RNA polymerase; DNA sequencing; ss.

OS Chimeric - Escherichia coli.

OS Chimeric - Bacteriophage T7.

OS Chimeric - Bacteriophage SP6.

XX Chimeric - Bacteriophage SP6.

PH Key

FT repeat_unit

FT 1..19

FT /tag= a

FT /rpt_type= INVERTED

FT /note= "O-end"

FT 3159..3177

FT /tag= b

FT /rpt_type= INVERTED

FT /note= "O-end"

FT complement (28..1)

FT /tag= c

FT /note= "Phage T7 promoter"

FT 3143..3177

FT /tag= d

FT /note= "Phage SP6 promoter"

FT 407..1201

FT /tag= e

FT /phenotype= kanamycin resistance

FT complement (2980..1549)

FT /tag= f

FT /product= transposase

FT

```
XX USS137829-A.
XX 11-AUG-1992.
XX 05-OCT-1987; 87US-0105422.
XX 05-OCT-1987; 87US-0105422.
XX (UNIW ) UNIV WASHINGTON.
XX Berg DB, Huang HV, Nag DK;
XX WPI; 1992-292374/35.
XX New DNA transposon derived from transposon Tn5 - useful for
XX generation of mutants and rapid screening of DNA sequences
XX Claim 3; Fig 4; 19pp; English.
XX Transposon Tn5seq1 was constructed from the pBR322-derived plasmid
XX pBR322::Tn5. The novel transposon is contained in plasmid pBRG1408. It
XX can be used for sequencing bacterial genomic DNA without recombinant
XX DNA cloning and also for sequencing DNA cloned in phages or
XX multicopy plasmids. It may also be useful for RNA sequencing with
XX SP6 and T7 RNA polymerases, for making RNA transcripts in vitro and
XX for hyperexpression or specific transcription of adjacent genes in
XX vivo.
XX (Updated on 10-MAR-2003 to add missing OS field.)
XX Sequence 3177 BP; 662 A; 898 C; 893 G; 724 T; 0 other;
XX
```

```
Query Match 100.0%; Score 1001; DB 13; Length 3177;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAACGATTCCGAAAGCCCAACCTTTTCATAG 60
Db 1340 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAACGATTCCGAAAGCCCAACCTTTTCATAG 1281
QY 61 AAGCGCGCGGTGGAATCGAAATCTGATGAGGAGGATAGAGAGGCGATCGCGT 120
Db 1280 AAGCGCGCGGTGGAATCGAAATCTGATGAGGAGGATAGAGAGGCGATCGCGT 1221
QY 121 TCGAAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGAGGCGATCGCT 180
Db 1220 TCGAAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGAGGCGATCGCT 1161
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAAGAGGAGCGGTAGCGCCCAATTCGCCCCAA 240
Db 1160 GCGAATCGGAGCGCGATACCGTAAAGCAAGAGGAGCGGTAGCGCCCAATTCGCCCCAA 1101
QY 241 GCTCTTCAGCAATATCAGCGGTAGCAAGCTATGCTCTGCTGATGAGGCGTCCGACACCA 300
Db 1100 GCTCTTCAGCAATATCAGCGGTAGCAAGCTATGCTCTGCTGATGAGGCGTCCGACACCA 1041
QY 301 GCCGCGCACAGTCCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 360
Db 1040 GCCGCGCACAGTCCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 981
QY 361 AGGCATCGCATGGGTACAGAGATCTCTGCGGTGCGGCGATGCGCGCTTGGCCTGG 420
Db 980 AGGCATCGCATGGGTACAGAGATCTCTGCGGTGCGGCGATGCGCGCTTGGCCTGG 921
QY 421 CGAAACAGTTTCGGCTGCGCGGAGCCCTGATGCTCTGCTGCTGATGATCATCTGATCCACAA 480
Db 920 CGAAACAGTTTCGGCTGCGCGGAGCCCTGATGCTCTGCTGCTGATGATCATCTGATCCACAA 861
QY 481 GACCGCTTCCATCCGAGTACGTGCTGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
Db 860 GACCGCTTCCATCCGAGTACGTGCTGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 801
QY 541 GGCAGGTAGCGGATCAAGCGGTATGCGCGCGCGGATTCGATCAGCCATGATGATCTT 600
```

Db 800 GGCAGGTAGCCGATCAAGGTATGACGCGCGCATTCGCATCAGCCATGATGATCTT 741
 Qy 601 TCTCGGACGAGCAAGGTAGATGACAGGAGATCTCTGCGCGCACTTCGCGCAATAGCA 660
 Db 740 TCTCGGACGAGCAAGGTAGATGACAGGAGATCTCTGCGCGCACTTCGCGCAATAGCA 681
 Qy 661 GCGAGTCCCTTCGCGTTCAGTCAACAGTTCGAGCAAGTTCGCGCAAGTTCGCGTTCG 720
 Db 680 GCGAGTCCCTTCGCGTTCAGTCAACAGTTCGAGCAAGTTCGCGCAAGTTCGCGTTCG 621
 Qy 721 TGCCACGACGATAGCGCGTTCGCTCTGCGCAAGTTCGCGCAAGTTCGCGCAAGTTCG 780
 Db 620 TGCCACGACGATAGCGCGTTCGCTCTGCGCAAGTTCGCGCAAGTTCGCGCAAGTTCG 561
 Qy 781 CGTCTTGACAAAGAACCGCGCGCTGCGCTGACAGCGCGCAAGTTCGCGCAAGTTCGCG 840
 Db 560 CGTCTTGACAAAGAACCGCGCGCTGCGCTGACAGCGCGCAAGTTCGCGCAAGTTCGCG 501
 Qy 841 AGCAGCGGATGTCTGTGCGCGCTGCGCTGACAGCGCGCAAGTTCGCGCAAGTTCGCG 900
 Db 500 AGCAGCGGATGTCTGTGCGCGCTGCGCTGACAGCGCGCAAGTTCGCGCAAGTTCGCG 441
 Qy 901 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAACGATCTCATCTGCTCTT 960
 Db 440 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAACGATCTCATCTGCTCTT 381
 Qy 961 GATCAGATCTTGATCCCTGCGCGCATCAGATCTTGGCGGC 1001
 Db 380 GATCAGATCTTGATCCCTGCGCGCATCAGATCTTGGCGGC 340

RESULT 5

AA80959/c

ID AA80959 standard; DNA; 3740 BP.

XX AC AA80959;

XX 25-MAR-2003 (updated)

DT 12-OCT-1990 (first entry)

XX DE Plasmid pDML1 encoding neomycin resistance and lac repressor.

XX XX affinity peptide; metal-chelate affinity chromatography;

XX KW neomycin phosphotransferase (neo); lac repressor (lacI); pSM1.1; ss.

XX OS synthetic.

XX FT Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 357..1151

XX FT /*tag= a

XX FT /product=neomycinphosphotransferase

XX FT 1432..2611

XX FT /*tag= b

XX FT /product=lac repressor

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX

PS Disclosure; Page ?; pp; German.

XX

CC Plasmid encodes a fusion protein

CC consisting of 1 or 2 affinity peptides bonded directly or

CC indirectly to a biologically active protein (i.e. dihydrofolate

CC reductase). The presence of adjacent His residues in the affinity

CC peptides allows purification of the biologically active protein by

CC metal-chelate chromatography on nitrilotriacetic acid resins.

CC See also AAP0401-3 and AA80955-8 and AA80960-63.

CC (Updated on 25-MAR-2003 to correct PR field.)

XX

SQ Sequence 3740 BP; 813 A; 1035 C; 1056 G; 836 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 841 AGCAGCCGATCTCTGTTGTCGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 900
 Db 450 AGCAGCCGATCTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 391
 QY 901 GAGAACCTCGCGTGCATCTCTGTTCAATCATCGGAAACGATCCTCATCTCTGTTCTT 960
 Db 390 GAGAACCTCGCGTGCATCTCTGTTCAATCATCGGAAACGATCCTCATCTCTGTTCTT 331
 QY 961 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 1001
 Db 330 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 290

RESULT 6

AAN81154/c

ID AAN81154 standard; DNA; 3740 BP.

XX AC

XX AAN81154;

XX DT 25-MAR-2003 (updated)

XX DT 09-JAN-2003 (updated)

XX DT 12-NOV-1990 (first entry)

XX DT 12-NOV-1990 (first entry)

XX DE Plasmid pDWI.1.

XX KW Plasmid pDWI.1.

XX OS Plasmodium falciparum.

XX FH Key

XX FT misc_feature

XX FT Location/Qualifiers

XX FT 357..1148

XX FT /*tag= a

XX FT /note="DNA encoding neomycinophosphotransferase."

XX FT 1532..2611

XX FT /*tag= b

XX FT /note="DNA encoding lac repressor."

XX FT

XX PN EP283829-A.

XX XX

XX PD 28-SEP-1988.

XX PF 08-MAR-1988; 88EP-0103564.

XX PR 19-MAR-1987; 87GB-0006599.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Cert U, Gentz R, Takacs B;

XX XX

XX DR WPI; 1988-272339/39.

XX PT New immunogenic polypeptides - derived from Plasmodium falciparum 190 kD

XX PT surface antigen precursor, useful in vaccines, and encoding DNA sequences

XX PS Disclosure; Page ?; pp; German.

XX XX

XX CC Plasmid pDWI.1 contains regions encoding neomycinophosphotransferase and

XX CC the lac repressor.

XX CC (Updated on 09-JAN-2003 to add missing OS field.)

XX CC (Updated on 25-MAR-2003 to correct PA field.)

XX CC (Updated on 25-MAR-2003 to correct PI field.)

XX XX

XX SQ Sequence 3740 BP; 813 A; 1036 C; 1055 G; 836 T; 0 other;

XX

XX Query Match

XX Best Local Similarity

XX Matches 1001; Conservative

XX 100.0%; Score 1001; DB 9; Length 3740;

XX 100.0%; Pred. No. 0;

XX Mismatches 0; Indels 0; Gaps 0;

XX

XX QY 1 CGCTGAGATCATCCAGCCGCGTCCCGGAAACGATTCGGAAGCCACCTTTTCATAG 60

XX Db 1290 CGCTGAGGATCATCCAGCCGCGTCCCGGAAACGATTCGGAAGCCACCTTTTCATAG 1231

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 61 AAGCGCGCGTGAATCGAAATCTCTGTATGACAGGTTCCGGCTCGCTTGGTCGTCATT 120
 Db 1230 AAGCGCGCGTGAATCGAAATCTCTGTATGACAGGTTCCGGCTCGCTTGGTCGTCATT 1171
 QY 121 TCGAACCCCAAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGGCT 180
 Db 1170 TCGAACCCCAAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGGCT 1111
 QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGAGCGGTGAGCCCAATTCGCCGCCAA 240
 Db 1110 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGAGCGGTGAGCCCAATTCGCCGCCAA 1051
 QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGTATAGCGGTCCGCACACCA 300
 Db 1050 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGTATAGCGGTCCGCACACCA 991
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAAGCGCCATTTTCCACCATGATATTCGGAAGC 360
 Db 990 GCCGGCCACAGTCGATGAATCCAGAAAAGCGCCATTTTCCACCATGATATTCGGAAGC 931
 QY 361 AGSCATCGCATGGGTCAAGAGATCTCTCGCGTCCGGCATGCGCGCTTGGAGCTGG 420
 Db 930 AGSCATCGCATGGGTCAAGAGATCTCTCGCGTCCGGCATGCGCGCTTGGAGCTGG 871
 QY 421 CGAAACAGTTCCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCCACAA 480
 Db 870 CGAAACAGTTCCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 811
 QY 481 GACCGCTTCCATCCGAGTACGTCGCTCGATCGATGTTTCGCTTGGTTCGATG 540
 Db 810 GACCGCTTCCATCCGAGTACGTCGCTCGATCGATGTTTCGCTTGGTTCGATG 751
 QY 541 GGCAGTACCGCGATCAAGCGTATCGAGCGCGCGGATTCATCAGCCCATGATGATCTT 600
 Db 750 GGCAGTACCGCGATCAAGCGTATCGAGCGCGCGGATTCATCAGCCCATGATGATCTT 691
 QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGATCTCTGCCCCGCACTTCGCCCAATAGCA 660
 Db 690 TCTCGCAGGAGCAAGGTGAGATGACAGGATCTCTGCCCCGCACTTCGCCCAATAGCA 631
 QY 661 GCCAGTCCCTTCGCGTTTCAGTGACAAAGTCAGACAGCTCGGACAGCTCGGAGAACCGCGTCG 720
 Db 630 GCCAGTCCCTTCGCGTTTCAGTGACAAAGTCAGACAGCTCGGACAGCTCGGAGAACCGCGTCG 571
 QY 721 TGGCCAGCCACGATAGCCGCTGCTCTGCTCTGAGTTTCAATCAGGCGCACCGGACAGT 780
 Db 570 TGGCCAGCCACGATAGCCGCTGCTCTGCTCTGAGTTTCAATCAGGCGCACCGGACAGT 511
 QY 781 CGGTCTTGACAAAAAGAACCGCGCGCCCTGCGCTGACAGCGGAAACACCGCGGATCAG 840
 Db 510 CGGTCTTGACAAAAAGAACCGCGCGCCCTGCGCTGACAGCGGAAACACCGCGGATCAG 451
 QY 841 AGCAGCGGATTCCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 900
 Db 450 AGCAGCGGATTCCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 391
 QY 901 GAGAACCTCGCGTGCATCTCTGTTCAATCATCGGAAACGATCCTCATCTCTGTCCTT 960
 Db 390 GAGAACCTCGCGTGCATCTCTGTTCAATCATCGGAAACGATCCTCATCTCTGTCCTT 331
 QY 961 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 1001
 Db 330 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 290

RESULT 7

AAN80483/c

ID AAN80483 standard; DNA; 3740 BP.

XX AC

XX AAN80483;

XX DT 25-MAR-2003 (updated)

XX DT 30-NOV-1990 (first entry)

XX

XX pDM1.1 contg. the neo-gene of Tn5 and the lacI-gene.
 XX HTLV-III; env protein; gag protein; AIDS; immunogenic activity;
 KW vaccines; diagnosis; oligonucleotides; pDM1.1; neo-gene; lacI-gene; ss.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT CDS 357..1148
 FT /*tag= a
 FT /product=neomycinphosphotransferase
 CDS 1532..2611
 FT /*tag= b
 FT /product=lac-repressor
 XX EP270114-A.
 XX 08-JUN-1988.
 XX 03-DEC-1987; 87EP-0117899.
 XX 05-DEC-1986; 86GB-0029116.
 XX (HOFF) HOFFMANN-LA ROCHE AG.
 XX Gentz R, Legrice S, Mous J, Stuber D;
 XX WPI; 1988-156343/23.
 XX New polypeptide with immunogenic activity of HTLV 3 env-gag protein
 FT - useful in vaccines and for AIDS diagnosis.
 XX Disclosure; Page ?; 49pp; German.
 XX Transposon Tn5: E. Beck et al., Gene 19, 327-336 (1982) for
 CC kanamycin resistance of E.coli.
 CC LacI-gene: P.J. Farabaugh, Nature 274, 765-769 (1978) contg. the
 CC mutation Iq (M.P. Calos, Nature 274, pp. 762-765 (1978) for the lac
 CC repressor.
 CC The plasmid also contains a region from plasmid pACYC184
 CC (A.C.Y. Chang and N.S. Cohen. J. Bacteriol. 134, 1141-1156 (1978) for
 CC stable replication.
 CC E.coli transfected with pDS5/RBSII.3A+5A (AAN80481), pDS8/RBSII
 CC (AAN80482) or derivatives, can'tg. a HTLV-gene, also contain plasmid
 CC pDM1.1. Expression is possible by addition of IPTG to the medium.
 CC gene is added.
 CC The DNA encodes peptides useful as active ingredients in vaccines
 CC and for detecting AIDS antibodies in serum etc.
 CC See also AAN80480-85.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 25-MAR-2003 to correct DR field.)
 XX Sequence 3740 BP; 813 A; 1040 C; 1051 G; 836 T; 0 other;
 XX Query Match 100.0%; Score 1001; DB 9; Length 3740;
 XX Best Local Similarity 100.0%; Pred. NO. 0;
 XX Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX 1 CGCTGGAGGATCATCCAGCGCGGTCGCCGGAAGACGATTCGAGCCCAACCTTTTCATAG 60
 DB 1290 CGCTGGAGGATCATCCAGCGCGGTCGCCGGAAGACGATTCGAGCCCAACCTTTTCATAG 1231
 XX 61 AAGCGCGGCGTGAATCGAAATCTCGTATGCGAGGTGGCGTCGCTTGGTCGTCATT 120
 DB 1230 AAGCGCGGCGTGAATCGAAATCTCGTATGCGAGGTGGCGTCGCTTGGTCGTCATT 1171
 XX 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCCCT 180
 DB 1170 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCCCT 1111

QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAACGAGGAGCGGTGAGCCCATTCGCCGCGCAA 240
 DB 1110 GCGAATCGGAGCGCGGATACCGTAAAGCAACGAGGAGCGGTGAGCCCATTCGCCGCGCAA 1051
 QY 241 GCTCTTCAGCAATATACGGGTAGCCAGCGTATGCTCTGATAGCGGTGCGGCACACCCA 300
 DB 1050 GCTCTTCAGCAATATACGGGTAGCCAGCGTATGCTCTGATAGCGGTGCGGCACACCCA 991
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
 DB 990 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 931
 QY 361 AGGCATCGCCATGGGTCAAGCAGATCTTCGCCGTGCGGCATGCGGCCTTGGAGCTGTG 420
 DB 930 AGGCATCGCCATGGGTCAAGCAGATCTTCGCCGTGCGGCATGCGGCCTTGGAGCTGTG 871
 QY 421 CGAACAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
 DB 870 CGAACAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 811
 QY 481 GACCGCTTCATCCCGAGTACGTGCTGCTCGATGCGATGTTTTCGTTGGTGTGCAATG 540
 DB 810 GACCGCTTCATCCCGAGTACGTGCTGCTCGATGCGATGTTTTCGTTGGTGTGCAATG 751
 QY 541 GGCAGGTAGCCGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGATGATCTT 600
 DB 750 GGCAGGTAGCCGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGATGATCTT 691
 QY 601 TCTCGCAGGACCAAGGTGAGATGACAGGAGATCTTCGCCCGGCATTCGCCCAATAGCA 660
 DB 690 TCTCGCAGGACCAAGGTGAGATGACAGGAGATCTTCGCCCGGCATTCGCCCAATAGCA 631
 QY 661 GCCAGTCCCTTCGCCGCTTCAGTACACGTCGAGCAGCTGCGCAAGGAAACGCCGCTCG 720
 DB 630 GCCAGTCCCTTCGCCGCTTCAGTACACGTCGAGCAGCTGCGCAAGGAAACGCCGCTCG 571
 QY 721 TGSCCAGCCACGATAGCGCGCTCGCTTCGTCGAGTTCATTCAGGCGCACCGGACAGGT 780
 DB 570 TGSCCAGCCACGATAGCGCGCTCGCTTCGTCGAGTTCATTCAGGCGCACCGGACAGGT 511
 QY 781 CGCTCTTGACAAAAGAACCGCGCGCCCTGCGCTGACAGCGGAAACACGGCGGCATCAG 840
 DB 510 CGCTCTTGACAAAAGAACCGCGCGCCCTGCGCTGACAGCGGAAACACGGCGGCATCAG 451
 QY 841 AGCAGCGGATGCTGTGTGTCGCGAGTCAATAGCGGATAGCCTCTCCACCCAGCGCGCG 900
 DB 450 AGCAGCGGATGCTGTGTGTCGCGAGTCAATAGCGGATAGCCTCTCCACCCAGCGCGCG 391
 QY 901 GAGAACCTGCGTCAATTCATCTTGTTCATATGCGAAACGATTCCTCATCTCTGTCTCTT 960
 DB 390 GAGAACCTGCGTCAATTCATCTTGTTCATATGCGAAACGATTCCTCATCTCTGTCTCTT 331
 QY 961 GATCAGATCTTGTATCCCTGCGGCATCAGATCTTCGCGGC 1001
 DB 330 GATCAGATCTTGTATCCCTGCGGCATCAGATCTTCGCGGC 290

RESULT 8
 AAQ06307/C
 ID AAQ06307 standard; DNA; 3740 BP.
 XX AAQ06307;
 AC AAQ06307;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-JAN-1991 (first entry)
 XX
 DE
 XX
 XX Sequence of plasmid pDM1.1.
 KW IFN-gamma receptor; autoimmune disease; multiple sclerosis;
 KW hypersensitivity; ds.
 XX
 OS Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT CDS 357...1146
 FT /*tag= a
 FT /label= Neomycin phosphotransferase.
 FT complement (1531...2609)
 FT /*tag= b
 FT /label= lac repressor.
 XX
 PN EP393502-A.
 XX
 XX 24-OCT-1990.
 XX
 XX 11-APR-1990; 90EP-0106992.
 XX
 XX 11-APR-1990; 90EP-0106992.
 XX
 PA (HOFF) HOFFMANN-LA ROCHE AG.
 XX
 PI Fountoulakis M, Garotta G, Stuber D;
 XX
 XX WPI; 1990-322042/43.
 DR
 XX Soluble interferon-gamma receptors - for treating auto-immune
 PT diseases, chronic inflammations, etc.
 PT
 XX Disclosure; Fig 16; 174pp; English.
 XX
 CC IFN-gamma is a therapeutically active agent in the treatment
 CC of autoimmune disease, allograft transplant rejections, multiple
 CC sclerosis, chronic inflammations and delayed hypersensitivity. It is
 CC also useful in identifying IFN-gamma agonists and antagonists.
 CC See also AAQ06301.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 XX Sequence 3740 BP; 813 A; 1035 C; 1056 G; 836 T; 0 other;
 SQ

Query Match 100.0%; Score 1001; DB 11; Length 3740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGTCCGGAAACAGATCCGAGCCCAACCTTTCATAG 60
 DB 1290 CGCTGGAGGATCATCCAGCGCGTCCGGAAACAGATCCGAGCCCAACCTTTCATAG 1231
 QY 61 AAGCGCGGCTGGAATCGAAATCTCGTATGCGAGGTGGGCGTGGCTTCGCTCGGTCAAT 120
 DB 1230 AAGCGCGGCTGGAATCGAAATCTCGTATGCGAGGTGGGCGTGGCTTCGCTCGGTCAAT 1171
 QY 121 TCGRACCCAGATCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATCGCT 180
 DB 1170 TCGRACCCAGATCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATCGCT 1111
 QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGAGGAAACGCTCAGCCCATTCGCCGCCAA 240
 DB 1110 GCGAATCGGAGCGCGATACCGTAAAGCAGAGGAAACGCTCAGCCCATTCGCCGCCAA 1051
 QY 241 GCTCTTACGATATACAGGTATGCCAAGCTATGCTCGATAGCGGTCCGCCACACCCA 300
 DB 1050 GCTCTTACGATATACAGGTATGCCAAGCTATGCTCGATAGCGGTCCGCCACACCCA 991
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAGC 360
 DB 990 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAGC 931
 QY 361 AGGCATCGCCATGGGTACGACGAGATCTTCGCGTCCGGCATCGCGCTTCGAGCTGG 420
 DB 930 AGGCATCGCCATGGGTACGACGAGATCTTCGCGTCCGGCATCGCGCTTCGAGCTGG 871
 QY 421 CGAACAGTTCGGCTGGCGGCGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
 DB 870 CGAACAGTTCGGCTGGCGGCGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 911
 QY 481 GACCGGCTTCATCCGAGTAGTGTCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 540

DB 810 GACCGGCTTCATCCAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 751
 QY 541 GCGAGGTAGCCGATCAAGCGTATGACGCGCGCGGATTCGATCAGCATGATGATGACTT 600
 DB 750 GCGAGGTAGCCGATCAAGCGTATGACGCGCGCGGATTCGATCAGCATGATGATGACTT 691
 QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCGCCCCGCGCACTTCGCCCAATAGCA 660
 DB 690 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCGCCCCGCGCACTTCGCCCAATAGCA 631
 QY 661 GCCAGTCCCTTCGCGCTTCAGTGACAACTCGAGCAGCAGTGGCGGAGGAAAGCCCGCTCG 720
 DB 630 GCCAGTCCCTTCGCGCTTCAGTGACAACTCGAGCAGCAGTGGCGGAGGAAAGCCCGCTCG 571
 QY 721 TGCCAGCCAGCATAGCCGCGCTGCTCTCGTCTCGAGTTCATTTCAGGGCAGCCGACAGGT 780
 DB 570 TGCCAGCCAGCATAGCCGCGCTGCTCTCGTCTCGAGTTCATTTCAGGGCAGCCGACAGGT 511
 QY 781 CGGTCTTGACAAAAGAACCGGCGCGCTCGCTGACAGCGGAAACACGCGGCGATCAG 840
 DB 510 CGGTCTTGACAAAAGAACCGGCGCGCTCGCTGACAGCGGAAACACGCGGCGATCAG 451
 QY 841 AGCAGCCGATTCGTCTGTTGTCGCCAGTCAATAGCCGAATAGCCTCTCCACCCCAAGCGCG 900
 DB 450 AGCAGCCGATTCGTCTGTTGTCGCCAGTCAATAGCCGAATAGCCTCTCCACCCCAAGCGCG 391
 QY 901 GAGAACCTCGGTGCAATTCATTTGTTCAATCATCGAAACGATCTCATCTGCTCTTT 960
 DB 390 GAGAACCTCGGTGCAATTCATTTGTTCAATCATCGAAACGATCTCATCTGCTCTTT 331
 QY 961 GATCAGATCTTCATCCCTCGCCATCAGATCCCTTGGCGGC 1001
 DB 330 GATCAGATCTTCATCCCTCGCCATCAGATCCCTTGGCGGC 290

RESULT 9
 AAQ12786/c
 ID AAQ12786 standard; DNA; 3740 BP.
 XX
 AC AAQ12786;
 XX
 DT 09-JAN-2003 (updated)
 DT 09-OCT-1991 (first entry)
 XX
 DE PDM1.1.
 XX
 KW Merozoite; antigen; vaccine; coccidiosis; RBS; ss.
 XX
 OS Unidentified.
 XX
 PN EP439056-A.
 XX
 PD 31-JUL-1991.
 XX
 PF 17-JAN-1991; 91EP-0100502.
 XX
 PR 26-JAN-1990; 90US-0470508.
 XX
 PA (HOFF) HOFFMANN-LA ROCHE AG.
 XX
 PI Binger MH;
 XX
 XX WPI; 1991-224487/31.
 DR
 XX Bimera merozoite surface antigen (23kD), DNA encoding it - and
 PT recombinant virus, useful in the prodn. of vaccine against
 PT coccidiosis in poultry
 XX
 PS Disclosure; Fig 11(a-c); 40pp; English.
 XX
 CC A 1.2 kb insert from lambda 5-7 was isolated, ends were repaired
 CC with Klenow polymerase and BamHI linkers were ligated to both ends.

CC The modified fragment was inserted into each of pDS56/RBSII,
CC pDS56/RBSII-1, and pDS56/RBSII-2. Plasmids contg. the inserts
CC in both possible orientations were transformed into E. coli
CC strain M15 carrying the compatible pDM1.1. The E. coli strain
CC M15 contg. plasmide pDS56/RBSII and pDM1.1 is described in
CC EP-316695.
CC See also AAQ12782-86.
CC (Updated on 09-JAN-2003 to add missing OS field.)
XX

QQ Sequence 3740 BP; 809 A; 1039 C; 1056 G; 836 T; 0 other;

Query Match 100.0%; Score 1001; DB 12; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGTGGAGGATCATCCAGCGCGGTCCTCCGGAAGAAACGATTCCGAAGCCCAACCTTTTCATAG	60
DB	1290	CGCTGGAGGATCATCCAGCGCGGTCCTCCGGAAGAAACGATTCCGAAGCCCAACCTTTTCATAG	1231
QY	61	AAGCGCGCGTGGAAATCGAAATCTCGTATGACGAGTTGGGCGTCTGGTTCGCTCAT	120
DB	1230	AAGCGCGCGTGGAAATCGAAATCTCGTATGACGAGTTGGGCGTCTGGTTCGCTCAT	1171
QY	121	TCGAACCCAGAGTCCCGCTCGAAGAACTCGTCAAGAGGCGATAGAGCGATGCGCT	180
DB	1170	TCGAACCCAGAGTCCCGCTCGAAGAACTCGTCAAGAGGCGATAGAGCGATGCGCT	1111
QY	181	GCGAATCGGAGCGGCGATACGTAAGACACAGAGAGCGGTTCAGCCCATTCGCGCGCAA	240
DB	1110	GCGAATCGGAGCGGCGGATACGTAAGACACAGAGAGCGGTTCAGCCCATTCGCGCGCAA	1051
QY	241	GCTCTTCAGCAATATCACGGGTAGCAACGCTATGCTGTATAGCGGTCCGCGCACCA	300
DB	1050	GCTCTTCAGCAATATCACGGGTAGCAACGCTATGCTGTATAGCGGTCCGCGCACCA	991
QY	301	GCGCGCCAGTCGATGATTCAGAAAGAGCGGCGATTTCCACCATGATATTCGCGCAAGC	360
DB	990	GCGCGCCAGTCGATGATTCAGAAAGAGCGGCGATTTCCACCATGATATTCGCGCAAGC	931
QY	361	AGGATCGCCATGGGTACAGAGATCTCGCGTGGGCGATGCGCGCTTGAGCGCTGG	420
DB	930	AGGATCGCCATGGGTACAGAGATCTCGCGTGGGCGATGCGCGCTTGAGCGCTGG	871
QY	421	CGAACAGTTCCGTCGCGGAGCCCTGATGCTTCTGTCAGATCATCTGATCGACAA	480
DB	870	CGAACAGTTCCGTCGCGGAGCCCTGATGCTTCTGTCAGATCATCTGATCGACAA	811
QY	481	GACCGGCTTCCATCCAGTACGTCTCGTTCGATCGATGTTTTCGCTTGGTTCGATG	540
DB	810	GACCGGCTTCCATCCAGTACGTCTCGTTCGATCGATGTTTTCGCTTGGTTCGATG	751
QY	541	GGCAGTAGCCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATCTT	600
DB	750	GGCAGTAGCCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATCTT	691
QY	601	TCTCGCAGGAGAGGTGATGATGACAGAGATCTCGCGCGCATTCGCGCGCATGACA	660
DB	690	TCTCGCAGGAGAGGTGATGATGACAGAGATCTCGCGCGCATTCGCGCGCATGACA	631
QY	661	GCCAGTCCCTTCCGCTTCAAGTACAACTCGACAGACAGTCCGCAAGGAGCGCCGTCG	720
DB	630	GCCAGTCCCTTCCGCTTCAAGTACAACTCGACAGACAGTCCGCAAGGAGCGCCGTCG	571
QY	721	TGCGCAGCAGATAGCGGCTGCTCTGTCGATGATTCATTCAGGCGACCGGACAGGT	780
DB	570	TGCGCAGCAGATAGCGGCTGCTCTGTCGATGATTCATTCAGGCGACCGGACAGGT	511
QY	781	CGGTCTTGACAAAAGAACCGGCGCGCTCGCTGACAGCGCGGAGCAAGCGGCGATCAG	840
DB	510	CGGTCTTGACAAAAGAACCGGCGCGCTCGCTGACAGCGCGGAGCAAGCGGCGATCAG	451
QY	841	AGCAGCGGATGTCTGTGTGTCAGTCTAGCCGATAGCTCTTCCACCCCAAGCGCGC	900

DB	450	AGCAGCGGATGTCTGTGTGTCGCGCATGATGCGGATAGCTCTCCACCAAGCGCGC	391
QY	901	GAGAACCTCGTGCATCATCTGTTCAATCATGCGAAGCATCTCATCCCTGCTCTT	960
DB	390	GAGAACCTCGTGCATCATCTGTTCAATCATGCGAAGCATCTCATCCCTGCTCTT	331
QY	961	GATCAGATCTTGATCCCTGCGCCATCAGATCTTGGCGGC	1001
DB	330	GATCAGATCTTGATCCCTGCGCCATCAGATCTTGGCGGC	290

RESULT 10
AAQ34610/C

ID AAQ34610 standard; DNA; 3740 BP.

XX AAQ34610;

XX AC

XX 25-MAR-2003 (updated)

DT 13-MAY-1993 (first entry)

XX XX

DE Plasmid pDM1.1.

XX Plasmid; pDS56/RBSII; -1; -2; regulatable; promoter; operator;

KW N25OPN25OP29; ribosomal binding site; RBS; RBSII(-1);

KW RBSII(-2); Pq25; E. coli; phage; T5; lac; repressor; lacI; lacIQ;

KW pDM1.1; replication; beta-lactamase; terminator; lambda; CAT; T1;

KW chloramphenicol acetyltransferase; rrnB operon; reading frame;

KW neomycin phosphotransferase; Tn5; transposon; selection marker;

KW pACYC184; replication; transmissibility; ss.

XX Synthetic.

OS

XX EP522482-A2.

PN

XX 13-JAN-1993.

PD

XX 06-JUL-1992; 92EP-0111407.

PF

XX 12-JUL-1991; 91US-0729099.

PR

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

PA

XX Binger M, Pasamontes L;

PI

XX WPI; 1993-010336/02.

DR

XX Immunogenic specified polypeptide and fragments - capable of

PT inducing an immune response against eimeria parasites and free

PT from proteins produced by these parasites

PS Disclosure; Fig 11; 72pp; English.

XX This sequence represents the plasmid pDM1.1. This plasmid was used

CC in conjunction with the expression plasmids pDS56/RBSII, -1, and -2.

CC These plasmids contain the regulatable promoter/operator element

CC N25OPN25OP29 and the ribosomal binding sites (RBS; RBSII, RBSII(-1)

CC promoter Pq25 of the E. coli phage T5. These plasmids may be

CC maintained in E. coli only if the promoter/operator element is

CC repressed by the binding of a lac repressor to the operator. The lac

CC repressor is encoded by the lacI gene. N25OPN25OP29 can be repressed

CC efficiently only when a sufficient amount of repressor molecules are

CC present in the cells. Therefore the lacI allele, which contains a

CC promoter mutant responsible for increased expression of the repressor

CC gene was used. This lacI allele is present on the plasmid pDM1.1

CC pDM1.1 also contains the neomycin phosphotransferase gene from the

CC Tn5 transposon, which is used as a selection marker, and a region of

CC the plasmid pACYC184 which contains all the information required for

CC replication and stable transmission to daughter cells. The region of

CC these plasmids between bases 1-1120 contains the replication region

CC and the gene for beta-lactamase. The beta-lactamase gene is modified

CC by elimination of the HindIII and PstI cleavage sites. These plasmids

CC also contain cleavage sites for the enzymes SalI, PstI and HindIII,

CC the terminator of E. coli phage lambda, the promoter-free gene of
 CC chloramphenicol acetyltransferase (CAT) and the terminator T1 of the
 CC E. coli rmb operon. These plasmids differ by one nucleotide
 CC following the ATG start codon resulting in protein expression in all
 CC three potential reading frames.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 3740 BP; 813 A; 1035 C; 1056 G; 836 T; 0 other;

Query Match 100.0%; Score 1001; DB 14; Length 3740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGTCCCGGAAAGCATTCGAAAGCCCAACCTTTCATAG 60
 DB 1290 CCTGAGGATCATCCAGCGCGTCCCGGAAAGCATTCGAAAGCCCAACCTTTCATAG 1231

QY 61 AAGCGCGGTTGGAATCGAATCTCGTGATGCGAGTTGGCGTTCGTTGGTTCGTCATT 120
 DB 1230 AAGCGCGGTTGGAATCGAATCTCGTGATGCGAGTTGGCGTTCGTTGGTTCGTCATT 1171

QY 121 TCGAACCCAGAGTCCCTCGAAGAACTGTCAGAGCGGATAGAGCGATGCGCT 180
 DB 1170 TCGAACCCAGAGTCCCTCGAAGAACTGTCAGAGCGGATAGAGCGATGCGCT 1111

QY 181 CGAATCGGAGCGGATACCGTAAAGCAGAGGAGCGTTCAGCCATTGCGCGCAA 240
 DB 1110 CGAATCGGAGCGGATACCGTAAAGCAGAGGAGCGTTCAGCCATTGCGCGCAA 1051

QY 241 GCTTTCAGATATCAGGATGAGCAAGCTATGTCGTATAGCGGTTCGCCACACCCA 300
 DB 1050 GCTTTCAGATATCAGGATGAGCAAGCTATGTCGTATAGCGGTTCGCCACACCCA 991

QY 301 GCGGCCACAGTCGATGAATCCAGAAAGCGCCATTTCCACCATGATATTCGCGAAGC 360
 DB 990 GCGGCCACAGTCGATGAATCCAGAAAGCGCCATTTCCACCATGATATTCGCGAAGC 931

QY 361 AGGCATCGGATGCGTACGAGATCCTCGCGTTCGCGATGCGCGCTTCGAGCTCG 420
 DB 930 AGGCATCGGATGCGTACGAGATCCTCGCGTTCGCGATGCGCGCTTCGAGCTCG 871

QY 421 CGAAGTTCGGTTCGCGAGCGGCTGATCTTCGTCAGATCCTTCGTCGATCGCAAA 480
 DB 870 CGAAGTTCGGTTCGCGAGCGGCTGATCTTCGTCAGATCCTTCGTCGATCGCAAA 811

QY 481 GACCGGCTTCATCCGATGATGCTCGCTCGATGCGATGTTTCGCTTGGTTCGATG 540
 DB 810 GACCGGCTTCATCCGATGATGCTCGCTCGATGCGATGTTTCGCTTGGTTCGATG 751

QY 541 GCGAGTACCGGATCAAGCTATGCGCGCGGATTCGATCAGCGCATGATGATGATCTT 600
 DB 750 GCGAGTACCGGATCAAGCTATGCGCGCGGATTCGATCAGCGCATGATGATGATCTT 691

QY 601 TCTCGGAGGAGCAGGTGAGATGACAGAGATCTCGCGCGGACCTTCGCCCAATAGCA 660
 DB 690 TCTCGGAGGAGCAGGTGAGATGACAGAGATCTCGCGCGGACCTTCGCCCAATAGCA 631

QY 661 GCGAGTACCGGATCAAGCTATGCGCGCGGATTCGATCAGCGCATGATGATGATCTT 720
 DB 630 GCGAGTACCGGATCAAGCTATGCGCGCGGATTCGATCAGCGCATGATGATGATCTT 571

QY 721 TGGCCAGCAGATAGCGCGCTCGCTCGATGCTTCATTCAGGCGCACCGACAGGT 780
 DB 570 TGGCCAGCAGATAGCGCGCTCGCTCGATGCTTCATTCAGGCGCACCGACAGGT 511

QY 781 CGGTTCGACAAAAGAACCGGCGCGCTCGCTCGATGCGCGGAAACACGCGCGCATCAG 840
 DB 510 CGGTTCGACAAAAGAACCGGCGCGCTCGCTCGATGCGCGGAAACACGCGCGCATCAG 451

QY 841 AGCAGCCATGCTGTTGTCGCGCATGATGCGGATAGCTTCCACCCAGCGCGCG 900
 DB 450 AGCAGCCATGCTGTTGTCGCGCATGATGCGGATAGCTTCCACCCAGCGCGCG 391

QY 901 GAGAACCTCGTGCAATCCATCTTGTTCATCATCGAAACGATCCTCCTGCTCTT 960
 DB 390 GAGAACCTCGTGCAATCCATCTTGTTCATCATCGAAACGATCCTCCTGCTCTT 331

QY 961 GATCAGATCTTGATCCCTGCGCCATCAGATCCTTGGCGCG 1001
 DB 330 GATCAGATCTTGATCCCTGCGCCATCAGATCCTTGGCGCG 290

RESULT 11

AAV33721
 ID AAV33721 standard; DNA; 3974 BP.
 XX AAV33721;
 AC AAV33721;
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE Bacterial expression vector pHE4a.
 XX
 KW Thymus receptor tyrosine kinase; TRTK; LERK; Eph; colon carcinoma;
 KW testicular carcinoma; pancreatic carcinoma; lung adenocarcinoma;
 KW breast carcinoma; hepatocellular carcinoma; cancer;
 KW cell proliferation; diagnosis; therapy; prognosis; human;
 KW vector; pHE4a; ds.
 XX
 OS Chimeric - Escherichia coli.
 OS Chimeric - Bacteriophage T5.
 XX

Key Location/Qualifiers
 FT -35_signal 3889..3894
 FT /tag= a
 FT misc_signal 3895..3911
 FT /tag= b
 FT /function= operator 1
 FT -10_signal 3912..3917
 FT /tag= c
 FT misc_signal 3925..3943
 FT /tag= d
 FT /function= operator 2
 FT RBS 3957..3962
 FT /tag= e
 XX
 WO9844111-A1.
 XX
 PD 08-OCT-1998.
 XX
 PD 27-MAR-1998; 98WO-US06021.
 XX
 PD 28-MAR-1997; 97US-0042856.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Soppet DR;
 XX
 WIPI; 1998-542702/46.
 XX

New isolated thymus receptor tyrosine kinase - is used to develop products for the diagnosis and treatment of cancers and other diseases involving aberrant cell proliferation

Disclosure; Page 76-79; 101pp; English.

This is the nucleotide sequence of bacterial expression vector pHE4a. This vector includes a neomycin phosphotransferase gene as a selectable marker, an Escherichia coli origin of replication (oriC from pUC129), a phage T5 promoter sequence, 2 lac operators (see also AAV33722), a Shine-Dalgarno sequence, the lactose operon repressor gene (lacI) and a multiple cloning site linker region. Expression of operatively linked sequences located downstream from the lac operators may be induced by addition of a lac operon inducer such as IPTG. pHE4 vectors have been produced that include the coding sequence (see AAV33697) for novel human thymus receptor tyrosine kinase (TRTK, see AAV70525-26). These can be used for the

CC production of recombinant TRK polypeptides in bacterial host cells.
CC The invention provides TRK polynucleotides and polypeptides.
CC vectors, host cells and recombinant methods for producing these.
CC Also provided are diagnostic methods for detecting disease states
CC associated with the aberrant expression of TRK and therapeutic
CC methods for treating such disease states, especially cancer.
XX

SQ Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTCTATG 60
DB 2579 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTCTATG 2638

QY 61 AAGCGCGCGTGAATCGAAATCTCGTATGTCAGGTCGGGTGCGTGTGTCGTCATT 120
DB 2639 AAGCGCGCGTGAATCGAAATCTCGTATGTCAGGTCGGGTGCGTGTGTCGTCATT 2698

QY 121 TCGAACCCAGAGTCCCGTCTCAAGAACTCTGTAAGAGGCGATGAAGGCGATGCGCT 180
DB 2699 TCGAACCCAGAGTCCCGTCTCAAGAACTCTGTAAGAGGCGATGAAGGCGATGCGCT 2758

QY 181 GCGAATCGGAGCGCGATACCGTAAAGCACAGGAGCGGTACGCCCATTCGCCGCCAA 240
DB 2759 GCGAATCGGAGCGCGATACCGTAAAGCACAGGAGCGGTACGCCCATTCGCCGCCAA 2818

QY 241 GCTCTTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCCAACCCCA 300
DB 2819 GCTCTTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCCAACCCCA 2878

QY 301 GCGCGGCACAGTCGATTCAGAAAGAGCGGCATTTTCACCATGATATTCGCGCAAGC 360
DB 2879 GCGCGGCACAGTCGATTCAGAAAGAGCGGCATTTTCACCATGATATTCGCGCAAGC 2938

QY 361 AGGCATCGGCATCGGTCACGACGAGATCCTCGCGTCGGGCATGCGGCTTGACCTGG 420
DB 2939 AGGCATCGGCATCGGTCACGACGAGATCCTCGCGTCGGGCATGCGGCTTGACCTGG 2998

QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGCAAA 480
DB 2999 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGCAAA 3058

QY 481 GACCGGCTTCCATCCAGTACGTCGCTCGATGCGATGTTTCGCTTGCTGTCGAATG 540
DB 3059 GACCGGCTTCCATCCAGTACGTCGCTCGATGCGATGTTTCGCTTGCTGTCGAATG 3118

QY 541 GGCAGTAGCCGATCAAGCGTATGAGCGCGCATTCGATCAGCCATGATGATACCTT 600
DB 3119 GGCAGTAGCCGATCAAGCGTATGAGCGCGCATTCGATCAGCCATGATGATACCTT 3178

QY 601 TCTCGCAGGAGCAAGTGCAGTACAGAGATCCTGCGCGGCTTCGCGCATTCGCCATAGCA 660
DB 3179 TCTCGCAGGAGCAAGTGCAGTACAGAGATCCTGCGCGGCTTCGCGCATTCGCCATAGCA 3238

QY 661 GCAGTCCCTTCCGCTTCAGTGACAACTCGAGCAGCTGCGGAAGAAAGCGCCGTCG 720
DB 3239 GCAGTCCCTTCCGCTTTCAGTGACAACTCGAGCAGCTGCGGAAGAAAGCGCCGTCG 3298

QY 721 TGCCACGACGATAGCGCGTCTGCTCTGCGATTCATTCAGGACCGGACAGGT 780
DB 3299 TGCCACGACGATAGCGCGTCTGCTCTGCGATTCATTCAGGACCGGACAGGT 3358

QY 781 CGGCTTTCGAAAGAAAGCGCGGCTTCGCTGACAGCGGACAGCGGCGGATCAG 840
DB 3359 CGGCTTTCGAAAGAAAGCGCGGCTTCGCTGACAGCGGACAGCGGCGGATCAG 3418

QY 841 AGCAGCGGATTTGCTGTGTCCTCATAGCCGAATAGCTCTCCACCCCAAGCGCGC 900
DB 3419 AGCAGCGGATTTGCTGTGTCCTCATAGCCGAATAGCTCTCCACCCCAAGCGCGC 3478

QY 901 GAGAACCTGCTGCAATCCATCTTGTTCATCATCGAAACGATCCTCATCTCTCTT 960
DB 3479 GAGAACCTGCTGCAATCCATCTTGTTCATCATCGAAACGATCCTCATCTCTCTT 3538

QY 961 GATCAGATCTTGATCCCTGCGGCATCAGATCCTTGGCGGC 1001
DB 3539 GATCAGATCTTGATCCCTGCGGCATCAGATCCTTGGCGGC 3579

RESULT 12

AAV66372
ID AAV66372 standard; cDNA; 3974 BP.

XX AAV66372;

XX 18-JAN-1999 (first entry)

XX Vector pHE4a.

XX Cardiac and pancreatic protein; CAPP; muscle-derived growth factor;
KW human; pancreatitis; myocardial infarction; cardiomyopathy;
KW vector; pHE4a; ds.

XX Chimeric - Escherichia coli.

XX Chimeric - bacteriophage T5.

XX Chimeric - synthetic.

XX WO9844112-A1.

XX 08-OCT-1998.

XX 27-MAR-1998; 98WO-US06022.

XX 28-MAR-1997; 97US-0042855.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Soppet DR;

XX WPI; 1998-557111/47.

XX New isolated cardiac and pancreatic protein - used to develop
PT products for the diagnosis and treatment of e.g. pancreatitis or
PT abnormal hypertrophy of the heart

XX Disclosure; Page 73-75; 113pp; English.

XX This is the nucleotide sequence of plasmid pHE4a (ATCC 209645). It
CC includes a neomycin phosphotransferase gene as selectable marker,
CC an Escherichia coli origin of replication, a T5 phage promoter
CC sequence, 2 lac operator sequences, a Shine-Delgarno sequence,
CC the lactose operon repressor gene (lacIq) and a multiple cloning
CC site linker region. A nucleotide sequence (see AAV66372) coding for
CC novel human cardiac and pancreatic protein (CAPP, see AAV66373) of
CC be operatively linked to the promoter and operator (see AAV66373) of
CC pHE4a, and CAPP expressed in E. coli cells following IPTG
CC induction. CAPP polypeptides can modulate the differentiation and
CC proliferation of cells and tissue, both in vivo and ex vivo. The
CC products can be used in the diagnosis and treatment of pancreatitis
CC and conditions that cause abnormal hypertrophy of the heart, such
CC as hypertension, myocardial infarction, valve disease and
CC cardiomyopathy. The products can also be used in detection and
CC cell culturing.

SQ Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTCTATG 60
DB 2579 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTCTATG 2638

QY 61 AAGCGGCGGTGGAATCGAAATCTCTGATGCGAGGTTGGCGTGGCTTGGTTCGTCATT 120
DB |||||
QY 2639 AAGCGGCGGTGGAATCGAAATCTCTGATGCGAGGTTGGCGTGGCTTGGTTCGTCATT 2698
DB |||||
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCAGAGGCGATAGAGGCGATGCGCT 180
DB |||||
QY 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCAGAGGCGATAGAGGCGATGCGCT 2758
DB |||||
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTTCGCCGCAA 240
DB |||||
QY 2759 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTTCGCCGCAA 2818
DB |||||
QY 241 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTCTGATAGCGGTCCGCCACCCCA 300
DB |||||
QY 2819 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTCTGATAGCGGTCCGCCACCCCA 2878
DB |||||
QY 301 GCGGCCACAGTCGATGATTCAGAAAGCGGCCATTTCACCATGATATTCGGCAAGC 360
DB |||||
QY 2879 GCGGCCACAGTCGATGATTCAGAAAGCGGCCATTTCACCATGATATTCGGCAAGC 2938
DB |||||
QY 361 AGGCATCGCATGGGTACGAGAGATCTCGCTCGGCGATGCGCGCTTGAGCCTGG 420
DB |||||
QY 2939 AGGCATCGCATGGGTACGAGAGATCTCGCTCGGCGATGCGCGCTTGAGCCTGG 2998
DB |||||
QY 421 CGAAGAGTTCGGCTGCGCGAGCCCTGATGCTCTTCCTCAGATCATCTGATCGCAA 480
DB |||||
QY 2999 CGAAGAGTTCGGCTGCGCGAGCCCTGATGCTCTTCCTCAGATCATCTGATCGCAA 3058
DB |||||
QY 481 GACCGGTTCCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGTTGGTGGTCAAGT 540
DB |||||
QY 3059 GACCGGTTCCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGTTGGTGGTCAAGT 3118
DB |||||
QY 541 GCAGGTAGCGGATCAGCGTATGCGCGCGCATGTCATCGACCATGATGATGATCT 600
DB |||||
QY 3119 GCAGGTAGCGGATCAGCGTATGCGCGCGCATGTCATCGACCATGATGATGATCT 3178
DB |||||
QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGAGATCTCGCCCGGACATTCGCCCAATAGCA 660
DB |||||
QY 3179 TCTCGGAGGAGCAAGGTGAGATGACAGAGATCTCGCCCGGACATTCGCCCAATAGCA 3238
DB |||||
QY 661 GCAGTTCCTTCCCGCTTACGTGACACGTGACAGAGTCTCGCCCGGACATTCGCCCAATAGCA 720
DB |||||
QY 3239 GCAGTTCCTTCCCGCTTACGTGACACGTGACAGAGTCTCGCCCGGACATTCGCCCAATAGCA 3298
DB |||||
QY 721 TGGCAGGACCATGATAGCGCGCTGCTCTGCTGAGTTCATTGAGGSCACCGGACAGGT 780
DB |||||
QY 3299 TGGCAGGACCATGATAGCGCGCTGCTCTGCTGAGTTCATTGAGGSCACCGGACAGGT 3358
DB |||||
QY 781' CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGAGCAGCGGAAACACCGCGGCATCAG 840
DB |||||
QY 3359 CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGAGCAGCGGAAACACCGCGGCATCAG 3418
DB |||||
QY 841 AGCAGCGATTCCTGTTGTGCGCCAGTTCATAGCGGAATAGCTCTCCACCCACAGGCGCG 900
DB |||||
QY 3419 AGCAGCGATTCCTGTTGTGCGCCAGTTCATAGCGGAATAGCTCTCCACCCACAGGCGCG 3478
DB |||||
QY 901 GAGAACCTCGTGCATTCATCTTGTTCATCATSCGAAACGATCTCATCTGCTCTT 960
DB |||||
QY 3479 GAGAACCTCGTGCATTCATCTTGTTCATCATSCGAAACGATCTCATCTGCTCTT 3538
DB |||||
QY 961 GATCAGATTCGATCCCTCGCCATCAGATCTTGGCGGC 1001
DB |||||
QY 3539 GATCAGATTCGATCCCTCGCCATCAGATCTTGGCGGC 3579
DB |||||

RESULT 13

AAV56699

ID AAV56699 standard; cDNA; 3974 BP.

XX

AC AAV56699;

XX

DT 26-NOV-1998 (first entry)

XX Plasmid pHE4-5 cDNA.
DE
XX
KW ELL2; RNA polymerase II elongation factor; human; neoplastic disorder;
KW leukaemia; inhibitor; diagnosis; susceptibility; ds.
OS Synthetic.
XX
PN WO9837194-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US03177.
XX
PR 19-FEB-1997; 97US-0038447.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Conaway JW, Conaway RC, Duan DR, Shilatifard A;
XX
XX WPI; 1998-467557/40.
XX
XX ELL2 RNA polymerase II elongation factor - and antibodies; for the
PT prevention and treatment of neoplastic disorders, e.g. leukaemia
PT
XX
PS Disclosure; Page 61-63; 84pp; English.
XX
XX This sequence is plasmid pHE4-5 which is used in a method to isolate and
CC characterise a human RNA polymerase II elongation factor ELL2 which can
CC be used in the design of treatments for neoplastic disorders (especially
CC leukemias). Antagonists of the polypeptides or nucleic acids that
CC inhibit the expression of the protein, can be used to inhibit ELL2
CC activity in a patient. A method for diagnosing or determining a
CC susceptibility to neoplastic disorders comprises assaying ELL2 gene
CC expression level in mammalian cells or body fluid and comparing these
CC levels with a standard, such that an increase or decrease in levels
CC indicates an increased or decreased susceptibility. The gene expression
CC level is assayed using the antibody, or by detecting ELL2 mRNA levels.
CC The host cells of may be used to identify compounds which bind to an
CC ELL2 polypeptide.
XX
XX Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;
Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGAGGATCATCCAGCCGCGCTCCGGAAGAAAGATTCGGAAGCCCAACCTTTTCATAG 60
DB 2579 CGCTGAGGATCATCCAGCCGCGCTCCGGAAGAAAGATTCGGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGGCGGTGGAATCGAAATCTCTGATGCGAGGTTGGCGTGGCTTGGTTCGTCATT 120
DB 2639 AAGCGGCGGTGGAATCGAAATCTCTGATGCGAGGTTGGCGTGGCTTGGTTCGTCATT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCAGAGGCGATAGAGGCGATGCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCAGAGGCGATAGAGGCGATGCGCT 2758
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTTCGCCGCAA 240
DB 2759 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTTCGCCGCAA 2818
QY 241 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTCTGATAGCGGTCCGCCACCCCA 300
DB 2819 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTCTGATAGCGGTCCGCCACCCCA 2878
QY 301 GCGGCCACAGTCGATGATTCAGAAAGCGGCCATTTCACCATGATATTCGGCAAGC 360
DB 2879 GCGGCCACAGTCGATGATTCAGAAAGCGGCCATTTCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCATGGGTACGAGAGATCTCGCTCGGCGATGCGCGCTTGAGCCTGG 420

Db 2939 AGGCATCGCCATGGGTGACAGAGATCTCGCCGCTCGGCGATCGCGCCTTGAGCCTGG 2998
QY 421 CGAACAGTTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCCTGATCGCAA 480
Db 2999 CGAACAGTTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCCTGATCGCAA 3058
QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 3118
QY 541 GCGAGTACCGGATCAAGCGTATGACGCGCGGCGATTCGATCGACCATGATGATGAT 600
Db 3119 GGAGTACCGGATCAAGCGTATGACGCGCGGCGATTCGATCGACCATGATGATGAT 3178
QY 601 TCTCGCAGAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
Db 3179 TCTCGCAGAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3238
QY 661 GCGAGTTCCTTCGCTTCAGTGACACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
Db 3239 GCGAGTTCCTTCGCTTCAGTGACACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3298
QY 721 TGCCAGCAGCAGATAGCCGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 780
Db 3299 TGCCAGCAGCAGTACGCGGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3358
QY 781 CGGCTTTGACAAAGAAACCGGCGCCCTGCGCTGACACAGTGAAGTGAAGTGAAGTGAAG 840
Db 3359 CGGCTTTGACAAAGAAACCGGCGCCCTGCGCTGACACAGTGAAGTGAAGTGAAGTGAAG 3418
QY 841 AGCAGCGGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGCAGCGGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
QY 901 GAGAACCTCGTGCAATCCATTTGTTCAATCATGCGAAACGATCCTCATCTGCTCTT 960
Db 3479 GAGAACCTCGTGCAATCCATTTGTTCAATCATGCGAAACGATCCTCATCTGCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGCGGC 3579

RESULT 14

ID AAV57971
XX AAV57971 standard; cDNA; 3974 BP.

AC AAV57971;

XX AAV57971;
DT 19-NOV-1998 (first entry)

XX Vector pHE4-5 containing human MOGP.

XX Human; myelin oligodendrocyte glycoprotein-like protein; MOGP;
XX immunoglobulin superfamily; Ig; cancer; inflammation; diagnosis;
XX detection; multiple sclerosis; demyelinating disease; myelin; ds.

OS Homo sapiens.

OS Synthetic.

XX WO9833912-A1.

XX 06-AUG-1998.

XX 30-JAN-1998; 98WO-US01593.

XX 30-JAN-1997; 97US-0035445.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Chopra A, Gentz RL, Olsen HS, Ruben SM;

PI

DR WPI; 1998-437466/37.
XX Human myelin oligodendrocyte glycoprotein-like protein - used to
PT develop products for the diagnosis, prognosis and treatment of
PT cancers, inflammation or demyelinating disease such as multiple
PT sclerosis
XX
XX Disclosure; Page 63-65; 100pp; English.
XX
XX The present sequence represents pHE405 vector which contains human
CC myelin oligodendrocyte glycoprotein-like protein (MOGP). The presence
CC of MOGP on the surface of oligodendrocytes is associated with
CC oligodendrocyte maturation and with the production of myelin. Products
CC from the present invention can be used for the diagnosis, prognosis and
CC treatment of e.g. cancers, inflammation, or demyelinating disease such
CC as multiple sclerosis.
XX
XX Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCGGAACCGATTCGGAAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGCGCTCCGGAACCGATTCGGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGGTGAATCGAAATCTCTGATGCGAGGTTGGCGTCTGCTTGGTGGTCAAT 120
Db 2639 AAGCGCGGTGAATCGAAATCTCTGATGCGAGGTTGGCGTCTGCTTGGTGGTCAAT 2698
QY 121 TCGAACCCCAAGTCCCGCTCAGAACTCTGTCAGAAAGCGGATAGAGGCGGATCGCT 180
Db 2699 TCGAACCCCAAGTCCCGCTCAGAACTCTGTCAGAAAGCGGATAGAGGCGGATCGCT 2758
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTCAGCCCAATTCGCCGCCAA 240
Db 2759 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTCAGCCCAATTCGCCGCCAA 2818
QY 241 GCTCTTCAGCAATATCAGCGGTAGCCACGCTATGTCCTGATAGCGGTCCGCCACCCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCCACGCTATGTCCTGATAGCGGTCCGCCACCCA 2878
QY 301 GCGCGCCACAGTCAATGAATCCAGAAAGCGGCAATTTCCACCAATGATATTCGCAAGC 360
Db 2879 GCGCGCCACAGTCAATGAATCCAGAAAGCGGCAATTTCCACCAATGATATTCGCAAGC 2938
QY 361 AGGCATCGCCATCGGTCAAGCAGATCTCCCGCTCGGCGATCGCGGCTTGAGCCTGG 420
Db 2939 AGGCATCGCCATCGGTCAAGCAGATCTCCCGCTCGGCGATCGCGGCTTGAGCCTGG 2998
QY 421 CGAACAGTTCGGCTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGCTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 3118
QY 541 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGCATTCGATCAGCATGATGATGAT 600
Db 3119 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGCATTCGATCAGCATGATGATGAT 3178
QY 601 TCTCGCAGAGCAAGTGAATGAAGAGATTCCTGCCCCCGGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGCAGAGCAAGTGAATGAAGAGATTCCTGCCCCCGGCACTTCGCCCAATAGCA 3238
QY 661 GCGAGTCCCTTCGCTTCAGTGACACAGTTCGAGCAGTGGCGAAGAACGCCGCTCG 720
Db 3239 GCGAGTCCCTTCGCTTCAGTGACACAGTTCGAGCAGTGGCGAAGAACGCCGCTCG 3298
QY 721 TGCCAGCAGCAGATAGCCGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 780

Qy 841 AGCAGCCGATTGTCTGTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGGCG 900
|||
Db 3419 AGCAGCCGATTGTCTGTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGGCG 3478
|||
Qy 901 GAGAACCTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCTGTCTCTT 960
|||
Db 3479 GAGAACCTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCTGTCTCTT 3538
|||
Qy 961 GATCAGATCTTGATCCCGCTGCGCCATCAGATCCTTGCGGCG 1001
|||
Db 3539 GATCAGATCTTGATCCCGCTGCGCCATCAGATCCTTGCGGCG 3579
|||

Search completed: February 15, 2004, 22:40:31
Job time : 325.065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 10:22:55 ; Search time 73.6147 Seconds
(without alignments)
6001.851 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000
Perfect score: 1001
Sequence: 1 CGCTGGAGGATCATCCAGCGCGGTCCCGAAACGATTCCGAAACCTTTTCATAG 1001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCUTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	2220	2	US-08-864-224-1
2	1001	100.0	3176	6	Sequence 1, Appli Patent No. 5212080
3	1001	100.0	3974	3	US-09-026-343-33
4	1001	100.0	3974	3	US-09-042-105-16
5	1001	100.0	3974	3	US-09-044-856A-7
6	1001	100.0	3974	3	US-09-023-082A-147
7	1001	100.0	3974	3	US-09-044-855A-7
8	1001	100.0	3974	3	US-09-078-670-4
9	1001	100.0	3974	4	US-09-026-408-14
10	1001	100.0	3974	4	US-09-362-871-33
11	1001	100.0	3974	4	US-09-627-154-4
12	1001	100.0	3974	4	US-09-027-287-50
13	1001	100.0	3974	4	US-09-437-602-4
14	1001	100.0	3974	4	US-09-252-656B-50
15	1001	100.0	3984	3	US-09-044-798A-10
16	1001	100.0	3984	4	US-09-725-460A-10
17	1001	100.0	4208	3	US-09-044-796A-9
18	1001	100.0	4208	4	US-09-725-460A-9
19	1001	100.0	4208	4	US-09-689-693-37
20	1001	100.0	4256	3	US-08-995-156A-21
21	1001	100.0	4256	4	US-09-419-281-21
22	1001	100.0	4277	3	US-09-044-796A-11
23	1001	100.0	4277	4	US-09-725-460A-11
24	1001	100.0	4328	3	US-09-132-808-1
25	1001	100.0	4328	3	US-08-910-647-2
26	1001	100.0	4328	4	US-09-620-925-2
27	1001	100.0	4328	4	US-09-620-260-1

28	1001	100.0	4328	4	US-09-620-259-1
29	1001	100.0	4800	4	US-09-554-929-1
30	1001	100.0	4818	3	US-08-910-647-4
31	1001	100.0	4818	4	US-09-620-925-4
32	1001	100.0	5107	3	US-08-910-647-3
33	1001	100.0	5107	4	US-09-620-925-3
34	1001	100.0	6561	4	US-09-380-190A-30
35	1001	100.0	8068	4	US-09-301-593-27
36	1001	100.0	8068	4	US-09-301-593-35
37	1001	100.0	8349	4	US-09-186-002-16
38	979	97.8	979	8	US-09-380-190A-20
39	979	97.8	979	8	US-09-380-190A-18
40	979	97.8	979	8	US-09-380-190A-19
41	979	97.8	979	8	US-09-380-190A-17
42	979	97.8	979	8	US-09-380-190A-15
43	979	97.8	979	8	US-09-380-190A-22
44	979	97.8	979	8	US-09-380-190A-16
45	979	97.8	979	8	US-09-380-190A-16

ALIGNMENTS

RESULT 1
US-08-864-224-1
; Sequence 1, Application US/08864224
; Patent No. 5851808
; GENERAL INFORMATION:
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Liu, Qinghua
; TITLE OF INVENTION: Rapid Subcloning Using Site-Specific
; TITLE OF INVENTION: Recombination
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,224
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: BCM-02681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-864-224-1

Query Match 100.0%; Score 1001; DB 2; Length 2220;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Gaps 0;
Oy 1 CGCTGGAGGATCATCCAGCGCGGTCCCGAAACGATTCCGAAACCTTTTCATAG 60
Db 932 CGCTGGAGGATCATCCAGCGCGGTCCCGAAACGATTCCGAAACCTTTTCATAG 991


```
QY 61 AAGCGCGGCTGGAATCGAATCTCGTATGAGGAGGCTGGGGCTGCTGGTCTGGTCAATT 120
Db 992 AAGCGCGGCTGGAATCGAATCTCGTATGAGGAGGCTGGGGCTGCTGGTCTGGTCAATT 1051
QY 121 TCGAATCCCAAGATCCCGCTCAGAGAACTCGTCAAGAAAGGCGATAGAGGCGATGCGCT 180
Db 1052 TCGAATCCCAAGATCCCGCTCAGAGAACTCGTCAAGAAAGGCGATAGAGGCGATGCGCT 1111
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGGAGGCGTCAAGCCATTCGCGCCAA 240
Db 1112 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGGAGGCGTCAAGCCATTCGCGCCAA 1171
QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTCTGATAGCGGTCCGCCACACCA 300
Db 1172 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTCTGATAGCGGTCCGCCACACCA 1231
QY 301 GCGGCGCACAGTCGATGATATCCAGAAAGCGGCAATTTCCACCATGATATTCGCAAGC 360
Db 1232 GCGGCGCACAGTCGATGATATCCAGAAAGCGGCAATTTCCACCATGATATTCGCAAGC 1291
QY 361 AGGCATCGCATGGGTACGACAGATCTCGCGCTCGGCGCATGGCGCTTTGAGCCTGG 420
Db 1292 AGGCATCGCATGGGTACGACAGATCTCGCGCTCGGCGCATGGCGCTTTGAGCCTGG 1351
QY 421 CGAAGATTCGGCTGGCGGAGCGGCTGATGCTCTTCCTCAGATCATCTGATCGACAA 480
Db 1352 CGAAGATTCGGCTGGCGGAGCGGCTGATGCTCTTCCTCAGATCATCTGATCGACAA 1411
QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 540
Db 1412 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 1471
QY 541 GCGAGTACCGGATCAAGCGTATGCGCGCGGCTTGCATCGCCATGATGATATCTT 600
Db 1472 GCGAGTACCGGATCAAGCGTATGCGCGCGGCTTGCATCGCCATGATGATATCTT 1531
QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTGCGCGGCACTTCGCCCAATAGCA 660
Db 1532 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTGCGCGGCACTTCGCCCAATAGCA 1591
QY 661 GCGAGTCCCTTCGCTTCAAGTCAAGCAAGTGCAGCAAGTGCAGCAAGTGCAGCAAGTGCAG 720
Db 1592 GCGAGTCCCTTCGCTTCAAGTCAAGCAAGTGCAGCAAGTGCAGCAAGTGCAGCAAGTGCAG 1651
QY 721 TCGCAGGACAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 1652 TCGCAGGACAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1711
QY 781 CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 1712 CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1771
QY 841 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 1772 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
QY 901 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1832 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
QY 961 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGGCGC 1001
Db 1892 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGGCGC 1932
```

RESULT 2

5212080-1/c

; Patent No. 5212080

; APPLICANT: NAG, DILIP K.; HUANG, HENRY V.; BERG, DOUGLAS E.

; TITLE OF INVENTION: METHOD OF DNA SEQUENCING USING DNA

; TRANSDUCER TNSSEAL

; NUMBER OF SEQUENCES: 6

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/508,379
; FILING DATE: 12-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 105,422
; FILING DATE: 05-OCT-1987
; SEQ ID NO:1:
; LENGTH: 3176
5212080-1
```

```
Query Match 100.0%; Score 1001; DB 6; Length 3176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCAGCGCGCTCCGGAAGACGATTCGGAAGCCCAACCTTTTCATAG 60
Db 1340 CGCTGGAGGATCATCAGCGCGCTCCGGAAGACGATTCGGAAGCCCAACCTTTTCATAG 1281
QY 61 AAGCGCGGCTGGAATCGAATCTCGTATGAGGAGGCTGGGGCTGCTGGTCTGGTCAATT 120
Db 1280 AAGCGCGGCTGGAATCGAATCTCGTATGAGGAGGCTGGGGCTGCTGGTCTGGTCAATT 1221
QY 121 TCGAATCCCAAGATCCCGCTCAGAGAACTCGTCAAGAAAGGCGATAGAGGCGATGCGCT 180
Db 1220 TCGAATCCCAAGATCCCGCTCAGAGAACTCGTCAAGAAAGGCGATAGAGGCGATGCGCT 1161
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGGAGGCGTCAAGCCATTCGCGCCAA 240
Db 1160 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGGAGGCGTCAAGCCATTCGCGCCAA 1101
QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTCTGATAGCGGTTCGCCACACCA 300
Db 1100 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTCTGATAGCGGTTCGCCACACCA 1041
QY 301 GCGGCGCACAGTCGATGATATCCAGAAAGCGGCAATTTCCACCATGATATTCGGAAGC 360
Db 1040 GCGGCGCACAGTCGATGATATCCAGAAAGCGGCAATTTCCACCATGATATTCGGAAGC 981
QY 361 AGCAGTCCGATCGGTCAAGAGATCTCGCGCTCGGCGCATGCGCGCTTGGAGCTGG 420
Db 980 AGCAGTCCGATCGGTCAAGAGATCTCGCGCTCGGCGCATGCGCGCTTGGAGCTGG 921
QY 421 CGAAGATTCGCTTCAAGTCAAGCAAGTGCAGTTCGATGCTGCTGCTGCTGCTGCTGCTG 480
Db 920 CGAAGATTCGCTTCAAGTCAAGCAAGTGCAGTTCGATGCTGCTGCTGCTGCTGCTG 861
QY 481 GACCGGCTTCATCCGAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 860 GACCGGCTTCATCCGAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
QY 541 GCGAGTACCGGATCAAGCGTATGCGCGCGGCACTTCGATCAGCCATGATGATCTT 600
Db 800 GCGAGTACCGGATCAAGCGTATGCGCGCGGCACTTCGATCAGCCATGATGATCTT 741
QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCGCCCGGCACTTCGCCCAATAGCA 660
Db 740 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCGCCCGGCACTTCGCCCAATAGCA 681
QY 661 GCGAGTCCCTTCGCTTCAAGTCAAGCAAGTGCAGTTCGATGCTGCTGCTGCTGCTGCTG 720
Db 680 GCGAGTCCCTTCGCTTCAAGTCAAGCAAGTGCAGTTCGATGCTGCTGCTGCTGCTGCTG 621
QY 721 TCGCAGGACAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 620 TCGCAGGACAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
QY 781 CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 560 CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
QY 841 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 500 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
```

QY 901 GAGAACCTCGGTGCAATCACTTCTTCAATCATGCGAAACGATCCTCATCTCTCTT 960
DB 440 GAGAACCTCGGTGCAATCACTTCTTCAATCATGCGAAACGATCCTCATCTCTCTT 381
QY 961 GATCAGATCTTGATCCCTCGGCCATCAGATCCTTGGCGGC 1001
DB 380 GATCAGATCTTGATCCCTCGGCCATCAGATCCTTGGCGGC 340

RESULT 3

US-09-026-343-33
; Sequence 33, Application US/09026343
; Patent No. 6008018
; GENERAL INFORMATION:
; APPLICANT: DUAN, D. ROXANNE
; APPLICANT: SHILATIFARD, ALI
; APPLICANT: CONAWAY, JOAN W.
; APPLICANT: CONAWAY, RONALD C.
; TITLE OF INVENTION: ELL2, A New Member of an ELL Family of
; TITLE OF INVENTION: RNA Polymerase II Elongation Factors
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/038,447
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1488.0880001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 372-2540
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
US-09-026-343-33

Query Match 100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGGTCCGGAAACGATCCGAAGCCCAACCTTTTCATAG 60
DB 2579 CGCTGGAGGATCATCCAGCGCGGTCCGGAAACGATCCGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGGTGGAATCGAATCTCGTATGGCAGAGTTGGCGTCCGTTGGTCGTCATT 120
DB 2639 AAGCGCGGTGGAATCGAATCTCGTATGGCAGAGTTGGCGTCCGTTGGTCGTCATT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGGATAGAGGCGATCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGGATAGAGGCGATCGCT 2758

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAAGCCCATTCGCGCCAA 240
DB 2759 GCGAATCGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAAGCCCATTCGCGCCAA 2818
QY 241 GCTCTTTCAGCAATATACAGGCTAGCAACGCTATGCTCTGATAGCGGTCCGCGCACACCA 300
DB 2819 GCTCTTTCAGCAATATACAGGCTAGCAACGCTATGCTCTGATAGCGGTCCGCGCACACCA 2878
QY 301 GCGGCGCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTTTCGCAAGC 360
DB 2879 GCGGCGCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTTTCGCAAGC 2938
QY 361 AGGCATCGCCATGCGGTCAAGAGGATCTCGCGTCCGCGATCGCGGCTTTCGCGCTGG 420
DB 2939 AGGCATCGCCATGCGGTCAAGAGGATCTCGCGTCCGCGATCGCGGCTTTCGCGCTGG 2998
QY 421 CGAACAGTTTCGCTCGCGCGAGCCCTGATGCTCTTCTGCTCCAGATCATCTTCGACAAA 480
DB 2999 CGAACAGTTTCGCTCGCGCGAGCCCTGATGCTCTTCTGCTCCAGATCATCTTCGACAAA 3058
QY 481 GACCGCTTCCATCCAGTACGTGCTCGCTGATGCGATGTTTCGCTTGGTGGTCAATG 540
DB 3059 GACCGCTTCCATCCAGTACGTGCTCGCTGATGCGATGTTTCGCTTGGTGGTCAATG 3118
QY 541 GCGAGTAGCCGATCAAGCGGTATGACGCGCGCATTCGATGCGATGATGATGATCTT 600
DB 3119 GCGAGTAGCCGATCAAGCGGTATGACGCGCGCATTCGATGCGATGATGATGATCTT 3178
QY 601 TCTCGCGAGGAGCAAGGTGAGATGACGAGATCTTCCGCGGACACTTCGCGCCCAATAGCA 660
DB 3179 TCTCGCGAGGAGCAAGGTGAGATGACGAGATCTTCCGCGGACACTTCGCGCCCAATAGCA 3238
QY 661 GCGAGTCCCTTCCCGCTTCAGTGACACGTCGAGCAGACGTCGCGCAAGGACCGCGTCG 720
DB 3239 GCGAGTCCCTTCCCGCTTCAGTGACACGTCGAGCAGACGTCGCGCAAGGACCGCGTCG 3298
QY 721 TGGCAGCCAGATAGCGCGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 780
DB 3299 TGGCAGCCAGATAGCGCGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 3358
QY 781 CGGCTCTTCACAAAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 3359 CGGCTCTTCACAAAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
QY 841 AGCAGCCGATTTGCTGTTGTCGCCAGTATAGCCGAATAGCTCTCCACCCCAAGCGCGC 900
DB 3419 AGCAGCCGATTTGCTGTTGTCGCCAGTATAGCCGAATAGCTCTCCACCCCAAGCGCGC 3478
QY 901 GAGAACCTCGGTGCAATCCTATCTTGTTCATATCGGAAGGATCCTCATCTCTCTCTT 960
DB 3479 GAGAACCTCGGTGCAATCCTATCTTGTTCATATCGGAAGGATCCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGGCCATCAGATCCTTGGCGGC 1001
DB 3539 GATCAGATCTTGATCCCTCGGCCATCAGATCCTTGGCGGC 3579

RESULT 4

US-09-042-105-16
; Sequence 16, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA

```

;
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: DNA
; US-09-042-105-16

```

```

Query Match 100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAAGCATTCGGAAGCCCAACCTTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAAGCATTCGGAAGCCCAACCTTTCATAG 2638
QY 61 AAGGCGCGGTGGAATCGAAATCTCGTGATGACGTTGGGCTCGCTTGGTGGTCAAT 120
Db 2639 AAGGCGCGGTGGAATCGAAATCTCGTGATGACGTTGGGCTCGCTTGGTGGTCAAT 2698
QY 121 TGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGATAGAGCGGATGCGCT 180
Db 2699 TGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGATAGAGCGGATGCGCT 2758
QY 181 GCGATCGGAGCGGATACCGTAAAGCAGGAGCGGTACGCCATTCGGCGCAA 240
Db 2759 GCGATCGGAGCGGATACCGTAAAGCAGGAGCGGTACGCCATTCGGCGCAA 2818
QY 241 GCTTTACCAATATCAGGGTAGCCAAAGCTATGCTGTGATGAGCGGTCCGCCACACCA 300
Db 2819 GCTTTACCAATATCAGGGTAGCCAAAGCTATGCTGTGATGAGCGGTCCGCCACACCA 2878
QY 301 GCGGCGCAGTGGATGATCAGAAAGCGGCCATTTTCACACATGATATTCGGCAAGC 360
Db 2879 GCGGCGCAGTGGATGATCAGAAAGCGGCCATTTTCACACATGATATTCGGCAAGC 2938
QY 361 AGGATCGGATCGGTTCAGACGAGATCCTCGCCGTGGGATGCGCGCTTGGACCTGG 420
Db 2939 AGGATCGGATCGGTTCAGACGAGATCCTCGCCGTGGGATGCGCGCTTGGACCTGG 2998
QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058

```

```

RESULT 5
US-09-044-856A-7
; Sequence 7, Application US/09044856A
; Patent No. 6075124
; GENERAL INFORMATION:
; APPLICANT: LI, HAODONG
; APPLICANT: RUBEN, STEVEN M
; APPLICANT: SUTTON, GRANGER G III
; TITLE OF INVENTION: HUMAN CHEMOTACTIC PROTEIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/044,856A
; FILING DATE: HEREWITH
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/726,830
; FILING DATE: 08-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,126
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,425
; FILING DATE: 21-APR-1995

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05384
FILING DATE: 16-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0340006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2500
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3974 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
US-09-044-856A-7

Query Match 100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGATCATCCAGCGCGCTCCCGAAAGAGATTCGAAAGCCCAACCTTTTCATAG 60
DB 2579 |||||
QY 61 AAGGCGCGGTGGAATCTGATGCGAGGTTGGCGTCGCTTGTGCGTCATT 120
DB 2639 AAGGCGCGGTGGAATCTGATGCGAGGTTGGCGTCGCTTGTGCGTCATT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGTCTGTCAGAGCGGATAGAGCGGATGCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAAAGTCTGTCAGAGCGGATAGAGCGGATGCGCT 2758
QY 181 GCGAATCGGAGCGCGATACCTGTAAGCAGAGCGGATGCGGATTCGCCGCAA 240
DB 2759 GCGAATCGGAGCGCGATACCTGTAAGCAGAGCGGATGCGGATTCGCCGCAA 2818
QY 241 GCTCTTCAGCAATATCAGGATGCGCAACGCTATGTCGATAGCGTCCGCGCACACCCA 300
DB 2819 GCTCTTCAGCAATATCAGGATGCGCAACGCTATGTCGATAGCGTCCGCGCACACCCA 2878
QY 301 GCCGCGCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGCGAAGC 360
DB 2879 GCCGCGCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGCGAAGC 2938
QY 361 AGGCATCGCCATGGGTACAGCAGAGATCCTCGCGGTGGGATGCGGCGCTTCAGGCTGG 420
DB 2939 AGGCATCGCCATGGGTACAGCAGAGATCCTCGCGGTGGGATGCGGCGCTTCAGGCTGG 2998
QY 421 CGAACAGTTCGGTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATGCGACAA 480
DB 2999 CGAACAGTTCGGTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATGCGACAA 3058
QY 481 CACCGGCTTCATCCGAGTACGTCTGCTGATGCGATGTTTCGCTTGGTGGTGAATG 540
DB 3059 CACCGGCTTCATCCGAGTACGTCTGCTGATGCGATGTTTCGCTTGGTGGTGAATG 3118
QY 541 GCGAGGTAGCGGATCAAGGTATGCGAGCGCGCATTTGCATCAGCCATGATGATACIT 600
DB 3119 GCGAGGTAGCGGATCAAGGTATGCGAGCGCGCATTTGCATCAGCCATGATGATACIT 3178
QY 601 TCTCGGAGGAGAGGTGATGACAGAGGATCTCCCGGCACTTCGCCCAATAGCA 660
DB 3179 TCTCGGAGGAGAGGTGATGACAGAGGATCTCCCGGCACTTCGCCCAATAGCA 3238
QY 661 GCCAGTCCCTTCGCGCTTCAGTGAACAAGTCGAGCAGAGTGGCGAAGAACCCCGCTCG 720
DB 3239 GCCAGTCCCTTCGCGCTTCAGTGAACAAGTCGAGCAGAGTGGCGAAGAACCCCGCTCG 3298
QY 721 TGGCCAGCCAGATAGCGGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

Db 3299 TGCCAGCCACGATAGCCGCGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
QY 781 CGGTCTTGACAAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 3359 CGGTCTTGACAAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
QY 841 AGCAGCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGCAGCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
QY 901 GAGAACCTGCTGCAATCCATCTTGTTCATCATGCGAAACGATCTCTCATCTGCTGCTGCT 960
Db 3479 GAGAACCTGCTGCAATCCATCTTGTTCATCATGCGAAACGATCTCTCATCTGCTGCTGCT 3538
QY 961 GATCAGATCTTATGATCCCTGCGCATCAGATCTTGGCGGC 1001
Db 3539 GATCAGATCTTATGATCCCTGCGCATCAGATCTTGGCGGC 3579

RESULT 6
US-09-023-082A-147
Sequence 147, Application US/09023082A
Patent No. 6077692
GENERAL INFORMATION:
APPLICANT: RUBEN, STEVEN M.
APPLICANT: JIMENEZ, PABLO
APPLICANT: DUAN, D. ROXANNE
APPLICANT: RAMPY, MARK A.
APPLICANT: MENDRICK, DONNA
APPLICANT: ZHANG, JUN
APPLICANT: NI, JIAN
APPLICANT: MOORE, PAUL A.
APPLICANT: COLEMAN, TIMOTHY A.
APPLICANT: GRUBER, JOACHIM R.
APPLICANT: DILLON, PATRICK J.
APPLICANT: GENTZ, REINER L.
TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR-2
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023.082A
FILING DATE: 13-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01790
FILING DATE: 14-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/461,195
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/023,852
FILING DATE: 13-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,045
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/862,432
FILING DATE: 23-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,875
FILING DATE: 13-AUG-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/055,561
 FILING DATE: 13-AUG-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFPE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.0360008/EKS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 147:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3974 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cdna
 US-09-023-082A-147

Query Match	100.0%;	Score 1001;	DB 3;	Length 3974;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1001;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	CGCTGGAGGATCATCCAGCGCGGCTCCCGAAAAACGATTCGGAAGCCCAACCTTTCATAG	60
Db	2579	CGCTGGAGGATCATCCAGCGCGGCTCCCGAAAAACGATTCGGAAGCCCAACCTTTCATAG	2638
QY	61	AAGCGGCGGTGGAAATCGAAATCTCGTGATGCGCAGGTTGGGCGTGCCTTGGTCGGTCATT	120
Db	2639	AAGCGGCGGTGGAAATCGAAATCTCGTGATGCGCAGGTTGGGCGTGCCTTGGTCGGTCATT	2698
QY	121	TCGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGGCGATAGAAGGCGATGCGCT	180
Db	2699	TCGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGGCGATAGAAGGCGATGCGCT	2758
QY	181	GGCAATCCGGAGCGCGGATACCTGTAAGACAGAGGAAGCGGTACGCCATTGCGCGCAA	241
Db	2759	GGCAATCCGGAGCGCGGATACCTGTAAGACAGAGGAAGCGGTACGCCATTGCGCGCAA	2818
QY	241	GCCTTTACGCAATATCACGGGTAGCAACGCTATGTCCTGATAGCGGTCGCGCACACCCA	300
Db	2819	GCCTTTACGCAATATCACGGGTAGCAACGCTATGTCCTGATAGCGGTCGCGCACACCCA	2878
QY	301	GGCGGCCACACTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATTGATATTCGGCAAGC	360
Db	2879	GGCGGCCACACTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATTGATATTCGGCAAGC	2938
QY	361	AGGCATCGCCATGGGTACAGACAGAGATCTTCGCGGTGCGGCGATGCGCCCTTGAGCCTGG	420
Db	2939	AGGCATCGCCATGGGTACAGACAGAGATCTTCGCGGTGCGGCGATGCGCCCTTGAGCCTGG	2998
QY	421	CGAACAGTTCCGCTGGCGCGAGCCCTGATGCTCTTCGTCAGATCACTCTGATCGCAAA	480
Db	2999	CGAACAGTTCCGCTGGCGCGAGCCCTGATGCTCTTCGTCAGATCACTCTGATCGCAAA	3058
QY	481	GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTTCGAATG	540
Db	3059	GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTTCGAATG	3118
QY	541	GGCAGGTAGCCGATCAAGCGTATGACGCGCGCGCATTCGATCAGCCCATGATGATACTT	600
Db	3119	GGCAGGTAGCCGATCAAGCGTATGACGCGCGCGCATTCGATCAGCCCATGATGATACTT	3178
QY	601	TCTCGCAGGACGAAGGTGAGATGACAGAGATCCTGCCCGGCATCTCGCCCAATAGCA	660
Db	3179	TCTCGCAGGACGAAGGTGAGATGACAGAGATCCTGCCCGGCATCTCGCCCAATAGCA	3238
QY	661	GCCAGTCCCTTCCCGCTTCACTGACAACTGCGACACAGCTGCGCAAGAACGCCCGCTCG	720
Db	3239	GCCAGTCCCTTCCCGCTTCACTGACAACTGCGACACAGCTGCGCAAGAACGCCCGCTCG	3298
QY	721	TGGCCAGCCAGATAGCCCGCTGCTCTGTCAGTTTCATTGAGGGCACCGACAGGT	780
Db	3299	TGGCCAGCCAGATAGCCCGCTGCTCTGTCAGTTTCATTGAGGGCACCGACAGGT	3358

RESULT 7

US-09-044-855A-7
; Sequence 7, Application US/09044855A
; Patent No. 6100389
; GENERAL INFORMATION:
; APPLICANT: LI, HAODONG
; APPLICANT: RUBEN, STEVEN M
; APPLICANT: SUTTON, GRANGER G III
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A HUMAN CHEMOTACTIC
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA

ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/044,855A
 FILING DATE: HERewith
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/479,126
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/424,425
 FILING DATE: 21-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/05384
 FILING DATE: 16-MAY-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFEE, ERIC K
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.0340005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3974 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cdna
 US-09-044-855A-7

Query Match	100.0%	Score 1001;	DB 3;	Length 3974;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAACGATTCGAGCCCAACCTTTCATAG 60
2579 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAACGATTCGAGCCCAACCTTTCATAG 2638
61 AAGCGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGGTGCGTTCGCTCAT 120
2639 AAGCGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGGTGCGTTCGCTCAT 2698
121 TCGAACCCAGAGTCCCGCTCAGAACTCTCAAGAGGCGATAGAGGCGATGCGCT 180
2699 TCGAACCCAGAGTCCCGCTCAGAACTCTCAAGAGGCGATAGAGGCGATGCGCT 2758
181 GCGAATCGGAGCGGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCGCCAA 240
2759 GCGAATCGGAGCGGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCGCCAA 2818
241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACCA 300
2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACCA 2878
301 GCGGCGCAGAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 360
2879 GCGGCGCAGAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 2938
361 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGCTCGGCGATGCGCGCTTGGACCTGG 420
2939 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGCTCGGCGATGCGCGCTTGGACCTGG 2998
421 CGAAGAGTTCGCTGGCGGAGCGCTGATGCTCTTCTGATGCGATCATCTGATGCGAA 480
2999 CGAAGAGTTCGCTGGCGGAGCGCTGATGCTCTTCTGATGCGATCATCTGATGCGAA 3058
481 GACCGGCTTCATCCGAGTACGCTGCTCGCTCGATGCGATGTTTCTGTTGGTGGTGAATG 540
3059 GACCGGCTTCATCCGAGTACGCTGCTCGCTCGATGCGATGTTTCTGTTGGTGGTGAATG 3118
541 GCGAGGTAGCGGATCAGCGTATGCGCGCGCATTCGATGCGATGATGATGATCTT 600
3119 GCGAGGTAGCGGATCAGCGTATGCGCGCGCATTCGATGCGATGATGATGATCTT 3178
601 TCTCGGAGGAGCAAGGTGAGATGACAGAGATCTGCGCGGACATTCGCGGCAATAGCA 660
3179 TCTCGGAGGAGCAAGGTGAGATGACAGAGATCTGCGCGGACATTCGCGGCAATAGCA 3238
661 GCGATGCTCTCCCGCTTCAGTACACAGTTCGAGCAGCTGCGCAGAGAACGCGCGCTCG 720
3239 GCGATGCTCTCCCGCTTCAGTACACAGTTCGAGCAGCTGCGCAGAGAACGCGCGCTCG 3298
721 TGGCCAGCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
3299 TGGCCAGCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3358
781 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
3359 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
841 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
3419 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
901 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
3479 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3538
961 GATCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
3539 GATCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3579

RESULT 8

US-09-078-670-4

; Sequence 4, Application US/09078670

; Patent No. 6143498
; GENERAL INFORMATION:
; APPLICANT: OLSEN, HENRIK S.
; APPLICANT: RUBEN, STEVEN M.
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,670
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,415
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0930001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
; US-09-078-670-4

Query Match 100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAACGATTCGAGCCCAACCTTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAACGATTCGAGCCCAACCTTTCATAG 2638
QY 61 AAGCGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGGTGCGTTCGCTCAT 120
Db 2639 AAGCGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGGTGCGTTCGCTCAT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAACTCTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAACTCTCAAGAGGCGATAGAGGCGATGCGCT 2758
QY 181 GCGAATCGGAGCGGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCGCCAA 240
Db 2759 GCGAATCGGAGCGGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCGCCAA 2818
QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACCA 2878
QY 301 GCGGCGCAGAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 360
Db 2879 GCGGCGCAGAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 2938
QY 361 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGCTCGGCGATGCGCGCTTGGACCTGG 420
Db 2939 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGCTCGGCGATGCGCGCTTGGACCTGG 2998

QY	421	CGAA	CAGTTT	CGGCTG	CGCGAG	CCCTG	ATGCTCTT	CGTCCAGAT	CATCTG	ATCGCAA	480
Db	2999	CGAA	CAGTTT	CGGCTG	CGCGAG	CCCTG	ATGCTCTT	CGTCCAGAT	CATCTG	ATCGCAA	3058
QY	481	GACCGG	CTTCAT	CCGAGT	AGTGCT	CGCTCG	ATGCGATG	TTTTCGGT	TGGTGGT	CGAATG	540
Db	3059	GACCGG	CTTCAT	CCGAGT	AGTGCT	CGCTCG	ATGCGATG	TTTTCGGT	TGGTGGT	CGAATG	3118
QY	541	GGCAGT	TAGCCGGAT	CAAGCG	TATG	CAGCCG	CCGCGCAT	TTCATC	CAGCCAT	GATGGAT	CTT 600
Db	3119	GGCAGT	TAGCCGGAT	CAAGCG	TATG	CAGCCG	CCGCGCAT	TTCATC	CAGCCAT	GATGGAT	CTT 3178
QY	601	TCTCGG	CAGGAG	CAAGTG	AGATG	ACAGG	AGATCT	TGCCCCG	GGCATCTT	CGCCCCA	ATAGCA 660
Db	3179	TCTCGG	CAGGAG	CAAGTG	AGATG	ACAGG	AGATCT	TGCCCCG	GGCATCTT	CGCCCCA	ATAGCA 3238
QY	661	GCCAGT	TCCTT	CCCGCTT	CAGTG	ACACG	TCGAGG	CAACAGT	CGGCAAG	GGACGCG	CGCTCG 720
Db	3239	GCCAGT	TCCTT	CCCGCTT	CAGTG	ACACG	TCGAGG	CAACAGT	CGGCAAG	GGACGCG	CGCTCG 3298
QY	721	TGGCCAG	CCACG	ATAGCCG	CGCTG	CTCTG	CTGACAGT	TTCATT	CAGGGG	CACCGG	ACAGGT 780
Db	3299	TGGCCAG	CCACG	ATAGCCG	CGCTG	CTCTG	CTGACAGT	TTCATT	CAGGGG	CACCGG	ACAGGT 3358
QY	781	CGGTTT	GACAAA	AAGAA	CCGGG	CGCCCT	CGCGCT	GACAGC	CGGAAC	CAGCGCG	CATCAG 840
Db	3359	CGGTTT	GACAAA	AAGAA	CCGGG	CGCCCT	CGCGCT	GACAGC	CGGAAC	CAGCGCG	CATCAG 3418
QY	841	AGCAGC	CGATG	TCTG	TGTTGG	CCGAGT	CATAGC	CGGAATAG	CCCTCT	CCACCC	CAAGCGGCG 900
Db	3419	AGCAGC	CGATG	TCTG	TGTTGG	CCGAGT	CATAGC	CGGAATAG	CCCTCT	CCACCC	CAAGCGGCG 3478
QY	901	GAGAAC	CTT	CGCTG	CAATCC	ATCTT	GTGTTCA	ATGCGAA	ACGATC	CTCATCT	CTGCTCTT 960
Db	3479	GAGAAC	CTT	CGCTG	CAATCC	ATCTT	GTGTTCA	ATGCGAA	ACGATC	CTCATCT	CTGCTCTT 3538
QY	961	GATCAG	ATCTT	GAATCC	CCCTG	GGCAT	ACAGAT	CAAGAT	CAAGAT	CAAGAT	CAAGAT 1001
Db	3539	GATCAG	ATCTT	GAATCC	CCCTG	GGCAT	ACAGAT	CAAGAT	CAAGAT	CAAGAT	CAAGAT 3579

RESULT 9

US-09-026-408-14
Sequence 14, Application US/09026408
Patent No. 6303338
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
TITLE OF INVENTION: INHIBITOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/026,408
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/934,011
FILING DATE: 15-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,056
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:

Db 3359 CGGTCTTGACAAAAGAACCGGGCGCCCTGGCTGACAGCGCGGAACGCGGGGCATCAG 3418
QY 841 ACACGCCGATCTCTGTGTGCGCGAGTCATAGCCGAATAGCTCTCCACCCCAAGCGGCG 900
Db 3419 ACACGCCGATCTCTGTGTGCGCGAGTCATAGCCGAATAGCTCTCCACCCCAAGCGGCG 3478
QY 901 GAGAACCTCGCTGCAATCCATCTTGTTCATCATCGGAAAGATCCCTCATCTCTCTTT 960
Db 3479 GAGAACCTCGCTGCAATCCATCTTGTTCATCATCGGAAAGATCCCTCATCTCTCTTT 3538
QY 961 GATCAGATCTTGATCCCTGGCCATCAGATCCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTGGCCATCAGATCCCTTGGCGGC 3579

RESULT 10

US-09-362-871-33
; Sequence 33, Application US/09362871
; Patent No. 6379923
; GENERAL INFORMATION:
; APPLICANT: DUAN, D. ROXANNE
; APPLICANT: SHILATIFARD, ALI
; APPLICANT: CONAWAY, JOAN W.
; APPLICANT: CONAWAY, RONALD C.
; TITLE OF INVENTION: ELI2, A New Member of an ELL Family of
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/362,871
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,343
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1488.0880001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 372-2540
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; US-09-362-871-33

Query Match 100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGGGCTCCCGGAAACGATTCCGAAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGGGCTCCCGGAAACGATTCCGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGCGTGAATCGAATCTCTGTATGCGAGGTGGCGTGGCTTGGTGGTCAATT 120
Db 2639 AAGCGCGCGTGAATCGAATCTCTGTATGCGAGGTGGCGTGGCTTGGTGGTCAATT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 2758
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCGGCCAA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCGGCCAA 2818
QY 241 GCTCTTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGTAGTAGGGTCCGCGCACACCA 300
Db 2819 GCTCTTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGTAGTAGGGTCCGCGCACACCA 2878
QY 301 GCGCGCCACAGTCAATGAAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 360
Db 2879 GCGCGCCACAGTCAATGAAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 2938
QY 361 AGGCATCGCATGGGTCAAGAGAGATCTCTGCGCTCGGGCATCGCGCTTGGCGCTGG 420
Db 2939 AGGCATCGCATGGGTCAAGAGAGATCTCTGCGCTCGGGCATCGCGCTTGGCGCTGG 2998
QY 421 CGAAGCTTGGCTCGCGCGAGCCCTCTGATGCTCTTCTCCAGATCATCTCTGATGACAA 480
Db 2999 CGAAGCTTGGCTCGCGCGAGCCCTCTGATGCTCTTCTCCAGATCATCTCTGATGACAA 3058
QY 481 GACCGCTTCCATCCGAGTACGTCTGCTCGATCGAGTGTTCGCTTGGTGGTCAATG 540
Db 3059 GACCGCTTCCATCCGAGTACGTCTGCTCGATCGAGTGTTCGCTTGGTGGTCAATG 3118
QY 541 GGCAGGTAGCGGATCAAGCGGTATGACCGCGCGCATTCGATCGCCATGATGATGATCTT 600
Db 3119 GGCAGGTAGCGGATCAAGCGGTATGACCGCGCGCATTCGATCGCCATGATGATGATCTT 3178
QY 601 TCTCGGAGGAGCAAGGTGAGATGAGAGAGATCTCTGCGCGCATCTTCGCGCCCAATAGCA 660
Db 3179 TCTCGGAGGAGCAAGGTGAGATGAGAGAGATCTCTGCGCGCATCTTCGCGCCCAATAGCA 3238
QY 661 GCCAGTCCCTTCCCGCTTCAGTGAACAAGTCGAGCAGAGTCCGCAAGGAGACCGCGTCG 720
Db 3239 GCCAGTCCCTTCCCGCTTCAGTGAACAAGTCGAGCAGAGTCCGCAAGGAGACCGCGTCG 3298
QY 721 TGGCAGCCAGATGATCGCGCTCGCTCTCTGCTCTCTGATTCATTGAGGCGACCGGACAGGT 780
Db 3299 TGGCAGCCAGATGATCGCGCTCGCTCTCTGCTCTCTGATTCATTGAGGCGACCGGACAGGT 3358
QY 781 CGGTCTTGACAAAAGAACCGGGCGCCCTCGCTGCGCTGACAGCCGGAACACGCGCGCATCAG 840
Db 3359 CGGTCTTGACAAAAGAACCGGGCGCCCTCGCTGCGCTGACAGCCGGAACACGCGCGCATCAG 3418
QY 841 AGCAGCCGATTTGCTGTTGTGCCAGTATAGCCGATAGCCCTCTCCACCCCAAGCGGCG 900
Db 3419 AGCAGCCGATTTGCTGTTGTGCCAGTATAGCCCTCTCTCCACCCCAAGCGGCG 3478
QY 901 GAGAACCTCGGTGCAATCCATCTTGTTCATCATCGAAGAGATCTCTCATCTCTCTCTT 960
Db 3479 GAGAACCTCGGTGCAATCCATCTTGTTCATCATCGAAGAGATCTCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGCGCATCAGATCCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCGCATCAGATCCCTTGGCGGC 3579

RESULT 11

US-09-627-154-4
; Sequence 4, Application US/09627154
; Patent No. 6420116
; GENERAL INFORMATION:
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Antimicrobial Peptide
; FILE REFERENCE: 1488.0930002
; CURRENT APPLICATION NUMBER: US/09/627,154
; CURRENT FILING DATE: 2000-07-27


```
; PRIOR APPLICATION NUMBER: US 09/078,670
; PRIOR FILING DATE: 1998-05-14
; PRIOR APPLICATION NUMBER: US 60/046,415
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: synthetic construct
US-09-627-154-4

Query Match      100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGATCATCCAGCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGGAGATCATCCAGCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 2638

QY 61 AAGCGCGCGTGGAAATCGAAATCTCGTGAAGGCGATAGAGGCGATGCGCT 180
Db 2639 AAGCGCGCGTGGAAATCTCGTGAAGGCGATAGAGGCGATGCGCT 2758

QY 121 TCGAACCCAGAGTCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 240
Db 2699 TCGAACCCAGAGTCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758

QY 181 GCGATCCGAGCGCGGATACCGTAAAGCAAGAGGCGGTCAGCCCATTCGCGCCAA 240
Db 2759 GCGATCCGAGCGCGGATACCGTAAAGCAAGAGGCGGTCAGCCCATTCGCGCCAA 2818

QY 241 GCTCTTCAGCAATATACGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCGCACACCA 300
Db 2819 GCTCTTCAGCAATATACGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCGCACACCA 300

QY 301 GCGCGCGCAGTCGATGTAATCCAGAAAGCGCGCATTTTCCACCATGATATTCGCGAAGC 360
Db 2879 GCGCGCGCAGTCGATGTAATCCAGAAAGCGCGCATTTTCCACCATGATATTCGCGAAGC 2938

QY 361 AGGCATCCGATGGGTACGACGATCTCGTCAAGAGGCGATAGAGGCGATGCGCT 420
Db 361 AGGCATCCGATGGGTACGACGATCTCGTCAAGAGGCGATAGAGGCGATGCGCT 420

; QUERY MATCH 100.0%; Score 1001; DB 4; Length 3974;
; BEST LOCAL SIMILARITY 100.0%; Pred. No. 0;
; MATCHES 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-09-027-287-50

Query Match      100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGATCATCCAGCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGGAGATCATCCAGCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 2638

QY 61 AAGCGCGCGTGGAAATCGAAATCTCGTGAAGGCGATAGAGGCGATGCGCT 120
Db 2639 AAGCGCGCGTGGAAATCTCGTGAAGGCGATAGAGGCGATGCGCT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758

QY 181 GCGATCCGAGCGCGGATACCGTAAAGCAAGAGGCGGTCAGCCCATTCGCGCCAA 240
Db 2759 GCGATCCGAGCGCGGATACCGTAAAGCAAGAGGCGGTCAGCCCATTCGCGCCAA 2818

QY 241 GCTCTTCAGCAATATACGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCGCACACCA 300
Db 2819 GCTCTTCAGCAATATACGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCGCACACCA 300

QY 301 GCGCGCGCAGTCGATGTAATCCAGAAAGCGCGCATTTTCCACCATGATATTCGCGAAGC 360
Db 2879 GCGCGCGCAGTCGATGTAATCCAGAAAGCGCGCATTTTCCACCATGATATTCGCGAAGC 2938

QY 361 AGGCATCCGATGGGTACGACGATCTCGTCAAGAGGCGATAGAGGCGATGCGCT 420
Db 361 AGGCATCCGATGGGTACGACGATCTCGTCAAGAGGCGATAGAGGCGATGCGCT 420
```

Db 2939 AGGCATCGCCATGGGTCAACGACGAGATCTTCGCGCTCGGCGATCGCGCGCTTTGACGCTGG 2998
QY 421 CGACAGCTTGGCTTGGGCGAGCCCTGATGCTCTTCGTCGAGATCATCTGATCGACAA 480
Db 2999 CGACAGTTCGGCTTGGGCGAGCCCTGATGCTCTTCGTCGAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCATCCGAGTACGTCGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCGAGTACGTCGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 3118
QY 541 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGATTCGATTCGATCGACCATGATGATCTT 600
Db 3119 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGATTCGATTCGATCGACCATGATGATCTT 3178
QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTTCGCGCGCGATTCGCGCGCGATGAC 660
Db 3179 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTTCGCGCGCGATTCGCGCGCGATGAC 3238
QY 661 GCGAGTCCCTTCGCGCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 720
Db 3239 GCGAGTCCCTTCGCGCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 3298
QY 721 TGGCCAGCCAGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 780
Db 3299 TGGCCAGCCAGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3358
QY 781 CGGTCTTGCACAAAGAACCGGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 840
Db 3359 CGGTCTTGCACAAAGAACCGGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3418
QY 841 AGCAGCGAGTGTCTGTGTGCGCCAGTTCATAGCGCGAGTTCATAGCGCGAGTTCATAG 900
Db 3419 AGCAGCGAGTGTCTGTGTGCGCCAGTTCATAGCGCGAGTTCATAGCGCGAGTTCATAG 3478
QY 901 GAGAACCTGGTGCAATCCATCTGTTCATGATGCGAAACGATTCCTCATCTCTCTCTT 960
Db 3479 GAGAACCTGGTGCAATCCATCTGTTCATGATGCGAAACGATTCCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTTCGCGCGCGATTCATGATCCCTTCGCGCGCG 1001
Db 3539 GATCAGATCTTGATCCCTTCGCGCGCGATTCATGATCCCTTCGCGCGCG 3579

RESULT 13

US-09-437-602-4

; Sequence 4, Application US/09437602

; Patent No. 6495128

; GENERAL INFORMATION:

; APPLICANT: Salcedo, Theodora W.

; APPLICANT: Patel, Vikram, P.

; APPLICANT: Nibbs, Robert J.B.

; APPLICANT: Graham, Gerard J.

; TITLE OF INVENTION: CHEMOKINE BETA-7

; FILE REFERENCE: 1488.033000F

; CURRENT APPLICATION NUMBER: US/09/437,602

; PRIOR FILING DATE: 1999-11-10

; PRIOR APPLICATION NUMBER: US 60/107,801

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 4

; LENGTH: 3974

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-437-602-4

Query Match

Best Local Similarity 100.0%; Score 1001; DB 4; Length 3974;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTCATAG 60

Db 2579 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTCATAG 2638

RESULT 14

US-09-252-656B-50

; Sequence 50, Application US/09252656B

; Patent No. 6495520

; GENERAL INFORMATION:

; APPLICANT: Ebner, Reinhard

; APPLICANT: Yu, Guo-Liang

APPLICANT: Ruben, Steven M.
APPLICANT: Zhang, Jun
APPLICANT: Ullrich, Stephen
APPLICANT: Zhai, Yifan
TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
FILE REFERENCE: 1488.0650006
CURRENT APPLICATION NUMBER: US/09/252,656B
CURRENT FILING DATE: 1999-02-19
PRIORITY FILING DATE: US 60/075,409
PRIORITY FILING DATE: 1998-02-20
PRIORITY FILING DATE: US 09/027,287
PRIORITY FILING DATE: 1998-02-20
PRIORITY FILING DATE: US 09/003,886
PRIORITY FILING DATE: 1998-01-07
PRIORITY FILING DATE: US 08/822,953
PRIORITY FILING DATE: 1997-03-21
PRIORITY FILING DATE: US 60/013,923
PRIORITY FILING DATE: 1996-03-22
PRIORITY FILING DATE: US 60/030,157
PRIORITY FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 50
LENGTH: 3974
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pHE4-5 expression vector
US-09-252-656B-50

Query Match 100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGAGAGCCCACTTTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGAGAGCCCACTTTTCATAG 2638

QY 61 AAGCGCGGGTGGAAATCTGTGATGCGAGGTTGGCGCTGCTTGTGCGGTCAAT 120
Db 2639 AAGCGCGGGTGGAAATCTGTGATGCGAGGTTGGCGCTGCTTGTGCGGTCAAT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAGACTGCTCAAGAGCGATAGAGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAGACTGCTCAAGAGCGATAGAGCGATGCGCT 2758

QY 181 GCGAATCGGGAGCGCGATACCGTAAAGACAGAGAGCGGTGAGCCATTCGCGGCAA 240
Db 2759 GCGAATCGGGAGCGCGATACCGTAAAGACAGAGAGCGGTGAGCCATTCGCGGCAA 2818

QY 241 GCTCTTCAGCAATATCAGGGTAGCCAAAGCGTATGCTGATAGCGGTTCGCGCAAC 300
Db 2819 GCTCTTCAGCAATATCAGGGTAGCCAAAGCGTATGCTGATAGCGGTTCGCGCAAC 2878

QY 301 GCGGCGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATCATATTCGGAAGC 360
Db 2879 GCGGCGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATCATATTCGGAAGC 2938

QY 361 AGGCAATCGGATGCGGTACGACGAGATCTCTCGCGCTCGGGCATCGCGCTTCGAGCCTG 420
Db 2939 AGGCAATCGGATGCGGTACGACGAGATCTCTCGCGCTCGGGCATCGCGCTTCGAGCCTG 2998

QY 421 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTGTCAGATCATCTGATCGACAA 3058

QY 481 GACCGGCTTCATCCGAGTACGTCTGCTCGATCGATGTTTTCGCTTGGTGGTTCGAATG 540
Db 3059 GACCGGCTTCATCCGAGTACGTCTGCTCGATCGATGTTTTCGCTTGGTGGTTCGAATG 3118

QY 541 GGCAGGTAGCGGATCAGCGTATGACGCGCGCCGCTTGCATCAGCCATGATGATCTT 600
Db 3119 GGCAGGTAGCGGATCAGCGTATGACGCGCGCCGCTTGCATCAGCCATGATGATCTT 3178

QY 601 TCTCGGAGGAGCAAGCTGAGATGACAGAGATCTTCCCGCGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGGAGGAGCAAGCTGAGATGACAGAGATCTTCCCGCGCACTTCGCCCAATAGCA 3238

QY 661 GCAGTCCCTTCCGCTTCAGTGACAACTGAGACAGCTGCGCAAGAAAGCGCGCTCG 720
Db 3239 GCAGTCCCTTCCGCTTCAGTGACAACTGAGACAGCTGCGCAAGAAAGCGCGCTCG 3298

QY 721 TGCCAGGACGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 3299 TGCCAGGACGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358

QY 781 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 3359 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418

QY 841 AGCAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGCAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478

QY 901 GAGAACTCGTGCAATCCATCTTGTTCATCATCGAAACGATCTCTGCTGCTT 960
Db 3479 GAGAACTCGTGCAATCCATCTTGTTCATCATCGAAACGATCTCTGCTGCTT 3538

QY 961 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 1001
Db 3539 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 3579

RESULT 15
US-09-044-796A-10
Sequence 10, Application US/09044796A
Patent No. 6194168
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner L.
APPLICANT: Coleman, Timothy A.
TITLE OF INVENTION: Expression Control Sequences
FILE REFERENCE: PV447
CURRENT APPLICATION NUMBER: US/09/044,796A
CURRENT FILING DATE: 1998-03-20
PRIORITY FILING DATE: 1997-09-30
PRIORITY FILING DATE: 09/044,796
PRIORITY FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 3984
TYPE: DNA
ORGANISM: Homo sapiens
US-09-044-796A-10

Query Match 100.0%; Score 1001; DB 3; Length 3984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGGAAGCCCACTTTTCATAG 60
Db 2701 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGGAAGCCCACTTTTCATAG 2760

QY 61 AAGCGCGGGTGGAAATCTGTGATGCGAGGTTGGCGCTGCTTGTGCGGTCAAT 120
Db 2761 AAGCGCGGGTGGAAATCTGTGATGCGAGGTTGGCGCTGCTTGTGCGGTCAAT 2820

QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGTCTGATGCGAGGTTGGCGCTGCTTGTGCGGTCAAT 180
Db 2821 TCGAACCCAGAGTCCCGCTCAGAAAGTCTGATGCGAGGTTGGCGCTGCTTGTGCGGTCAAT 2880

QY 181 GCGAATCGGGAGCGCGATACCGTAAAGACAGAGAGCGGTGAGCCATTCGCGGCAA 240
Db 2881 GCGAATCGGGAGCGCGATACCGTAAAGACAGAGAGCGGTGAGCCATTCGCGGCAA 2940

```
QY 241 GCTCTTTCAGCAATATCACGGGTAGCCAAAGCTATGTCTCTGATAGCGTCCGGCACACCCA 300
Db |||||||
QY 2941 GCTCTTTCAGCAATATCACGGGTAGCCAAAGCTATGTCTCTGATAGCGTCCGGCACACCCA 3000
Db |||||||
QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCACCATGATATTCGGCAAGC 360
Db |||||||
QY 3001 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCACCATGATATTCGGCAAGC 3060
Db |||||||
QY 361 AGGCATGCCATGGGTACAGGAGATCCTCGCGTGGGCGATGCGCGCTTGAGCCTGG 420
Db |||||||
QY 3061 AGGCATGCCATGGGTACAGGAGATCCTCGCGTGGGCGATGCGCGCTTGAGCCTGG 3120
Db |||||||
QY 421 CGAACAGTTCGGTGGCGGAGCCCTCGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db |||||||
QY 3121 CGAACAGTTCGGTGGCGGAGCCCTCGATGCTCTTCTGTCAGATCATCTGATCGACAA 3180
Db |||||||
QY 481 GACCGGCTTCATCGGAGTACGTGCTGCTCGATGCGATGTTTCGCTTGGTGGTGGCAATG 540
Db |||||||
QY 3181 GACCGGCTTCATCGGAGTACGTGCTGCTCGATGCGATGTTTCGCTTGGTGGTGGCAATG 3240
Db |||||||
QY 541 GGCAGGTAGCCGGATCAAGCGTATGCAGCGCGCGCATTTGATCAGCCATGATGATACTT 600
Db |||||||
QY 3241 GGCAGGTAGCCGGATCAAGCGTATGCAGCGCGCGCATTTGATCAGCCATGATGATACTT 3300
Db |||||||
QY 601 TCTCGGAGGAGCAAGGTGAGATCACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 660
Db |||||||
QY 3301 TCTCGGAGGAGCAAGGTGAGATCACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 3360
Db |||||||
QY 661 GCCAGTCCCTTCCCGCTTCAGTGACAGGTGCGAGCACAGCTGGCAGGAAACGCCCGTGG 720
Db |||||||
QY 3361 GCCAGTCCCTTCCCGCTTCAGTGACAGGTGCGAGCACAGCTGGCAGGAAACGCCCGTGG 3420
Db |||||||
QY 721 TGGCCAGCCAGATAGCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db |||||||
QY 3421 TGGCCAGCCAGATAGCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db |||||||
QY 781 CGGTCTTTCACAAAGAAACCGGGCGCCCTGCGCTGACAGCGGAAACAGCGGCAATCAG 840
Db |||||||
QY 3481 CGGTCTTTCACAAAGAAACCGGGCGCCCTGCGCTGACAGCGGAAACAGCGGCAATCAG 3540
Db |||||||
QY 841 AGCAGCCGATTTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db |||||||
QY 3541 AGCAGCCGATTTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
Db |||||||
QY 901 GAGAACCTGGTGAATCCATCTTGTTCATATCGAAGACGATCCTCATCTGCTCTT 960
Db |||||||
QY 3601 GAGAACCTGGTGAATCCATCTTGTTCATATCGAAGACGATCCTCATCTGCTCTT 3660
Db |||||||
QY 961 GATCAGATCTTGATCCCTGCGCATCAGATCCTTGGCGGC 1001
Db |||||||
QY 3661 GATCAGATCTTGATCCCTGCGCATCAGATCCTTGGCGGC 3701
```

Search completed: February 15, 2004, 19:10:18
Job time : 77.6147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 16:21:35 , Search time 420.484 Seconds
(without alignments)
8769.219 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 cgcggaggatcaccagcc.....gcatcagatccttgccgc 1001

Scoring table: OILGO_NUC

Gapex 60.0 , Gapex 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
C 1	1001	100.0	2406	13	US-10-307-138-14
C 2	1001	100.0	2772	13	Sequence 14, Appl
C 3	1001	100.0	3974	9	US-10-307-138-15
4	1001	100.0	3974	9	Sequence 15, Appl
5	1001	100.0	3974	9	Sequence 50, Appl
6	1001	100.0	3974	10	US-09-252-656B-50
7	1001	100.0	3974	10	US-09-972-912-8
8	1001	100.0	3974	11	US-09-903-684-14
9	1001	100.0	3974	11	US-09-935-726-16
10	1001	100.0	3974	11	US-09-105-470-25
11	1001	100.0	3974	11	US-09-345-373-147
12	1001	100.0	3974	12	US-09-373-658-12
13	1001	100.0	3974	12	US-09-983-687-12
14	1001	100.0	3974	12	US-10-375-680-50
15	1001	100.0	3974	13	US-10-614-275-25
	1001	100.0	3974	13	US-10-260-270-4
	1001	100.0	3974	13	US-10-197-844-9
					Sequence 9, Appli

Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 147, Appl
Sequence 16, Appl
Sequence 33, Appl
Sequence 7, Appl
Sequence 16, Appl
Sequence 147, Appl
Sequence 37, Appl
Sequence 21, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 4, Appl
Sequence 3, Appl
Sequence 9, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 353, Ap
Sequence 556, Ap
Sequence 13, Appl
Sequence 25, Appl
Sequence 27, Appl
Sequence 35, Appl
Sequence 16, Appl

US-10-120-398-16
US-10-120-377-16
US-10-120-414-16
US-10-035-212-147
US-09-921-143-16
US-10-028-780-33
US-10-141-965-7
US-10-084-488-16
US-10-075-446-147
US-10-261-950-37
US-10-054-967-21
US-10-307-138-4
US-10-445-642-1
US-10-278-751-2
US-10-322-360-1
US-10-278-751-4
US-10-307-138-3
US-10-278-751-3
US-09-897-006-9
US-09-921-143-12
US-09-921-143-36
US-09-764-891-5553
US-09-764-891-5586
US-10-307-138-13
US-10-090-983-1
US-10-054-536-25
US-10-159-006-27
US-10-159-006-35
US-10-198-478-16

ALIGNMENTS

RESULT 1

US-10-307-138-14/c
Sequence 14, Application US/10307138
Publication No. US20030162267A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroeder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
APPLICANT: Haefner, Stefan
APPLICANT: Kloppe, Corinna
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING REGULATORY
FILE REFERENCE: BGI-123CP2
CURRENT APPLICATION NUMBER: US/10307138
CURRENT FILING DATE: 2002-11-23
PRIOR APPLICATION NUMBER: 09/602874
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/422618
PRIOR FILING DATE: 2002-10-30
PRIOR APPLICATION NUMBER: 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142690
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: 60/151251
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19930476.9
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932122.1
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
Remaining Prior Application data removed - See File Wrapper or PALM.

```

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; OTHER INFORMATION: PINTTEGRATIV
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (457)..(1248)
; OTHER INFORMATION: kanamycin resistance gene from Tn5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1515)..(2375)
; OTHER INFORMATION: Ori from pMB for replication in E.coli
; US-10-307-138-14

Query Match      100.0%; Score 1001; DB 13; Length 2406;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGAGGATCATCCAGCGCGTCCCGGAAACGATTCGAAAGCCCAACCTTTTCATAG 60
DB 1390 CGCTGAGGATCATCCAGCGCGTCCCGGAAACGATTCGAAAGCCCAACCTTTTCATAG 1331
QY 61 AAGCGGGGGTGAATCGAATCTCGTATGCGAGGTGGCGGTGGCTTGGTGGTCAATT 120
DB 1330 AAGCGGGGGTGAATCGAATCTCGTATGCGAGGTGGCGGTGGCTTGGTGGTCAATT 1271
QY 121 TCGAACCCCGAGAGTCCCGCTCAGAGAACTCGTCAAGAACGCGATAGAGCGCATGCGCT 180
DB 1270 TCGAACCCCGAGAGTCCCGCTCAGAGAACTCGTCAAGAACGCGATAGAGCGCATGCGCT 1211
QY 181 GCGAATCGGAGCGCGATACGTAAGACGAGAGCGGTAGCGCATTCGCGGCAA 240
DB 1210 GCGAATCGGAGCGCGATACGTAAGACGAGAGCGGTAGCGCATTCGCGGCAA 1151
QY 241 GCTCTTCAGCAATATACGCGGTAGCAAGCTATGCTGATAGCGGTGCGGCACACCA 300
DB 1150 GCTCTTCAGCAATATACGCGGTAGCAAGCTATGCTGATAGCGGTGCGGCACACCA 1091
QY 301 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCCATTTCACCATGATATTCGCGAAGC 360
DB 1090 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCCATTTCACCATGATATTCGCGAAGC 1031
QY 361 AGGCATCGCATGGTCAAGCAGAGATCCTCGCGTGGGATCGCGGCTTGAAGCTGG 420
DB 1030 AGGCATCGCATGGTCAAGCAGAGATCCTCGCGTGGGATCGCGGCTTGAAGCTGG 971
QY 421 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTGTCTCAGATCATCTGATCGACAA 480
DB 970 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTGTCTCAGATCATCTGATCGACAA 911
QY 481 GACCGGCTTCATCCGAGTAGTGTCTGCTGATGCGATGTTGCTTGGTGGTGAATG 540
DB 910 GACCGGCTTCATCCGAGTAGTGTCTGCTGATGCGATGTTGCTTGGTGGTGAATG 851
QY 541 GCGAGGTAGCCGATCAAGCGTATGTCAGCGCGCATTTGATCATGAGCATGATGACTT 600
DB 850 GCGAGGTAGCCGATCAAGCGTATGTCAGCGCGCATTTGATCATGAGCATGATGACTT 791
QY 601 TCTCGGAGGAGCAAGTGAATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 660
DB 790 TCTCGGAGGAGCAAGTGAATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 731
QY 661 GCGATCCTTCCTCCGCTTCAGTCAACAGCTGACAGCATGTCGCAAGGAAACCGGTCG 720
DB 730 GCGATCCTTCCTCCGCTTCAGTCAACAGCTGACAGCATGTCGCAAGGAAACCGGTCG 671
QY 721 TGGCCAGCCAGATAGCGCGCTCGCTCGTCTGATTCATTCAGGGCACCGGACAGGT 780

```

670 TGCCAGCCAGCATAGCGCGCTCGTCTGTCAGTTTATTTCAGGGCACCGGACAGGT 611

781 CGGTCTTGACAAAGAACCGGGCGCCCTGCTGACAGCCGGAACACAGCGGGCATCAG 840

610 CGGTCTTGACAAAGAACCGGGCGCCCTGCTGACAGCCGGAACACAGCGGGCATCAG 551

841 AGCAGCCGATTTGCTCTTGTGTCCTCATAGCCGAAATAGCTCTCCACCAGCGCGCG 900

550 AGCAGCCGATTTGCTCTTGTGTCCTCATAGCCGAAATAGCTCTCCACCAGCGCGCG 491

901 GAGAACTCGGTGCAATCATCTTGTTCATCATGCGAAACGATCTCATCTGCTCTT 960

490 GAGAACTCGGTGCAATCATCTTGTTCATCATGCGAAACGATCTCATCTGCTCTT 431

961 GATCAGATCTTGATCCCTCGCGCATCAGATCCTTGGCGGC 1001

430 GATCAGATCTTGATCCCTCGCGCATCAGATCCTTGGCGGC 390

RESULT 2

US-10-307-138-15/c

; Sequence 15, Application US/10307138

; Publication No. US20030162267A1

; GENERAL INFORMATION:

; APPLICANT: Pompeius, Markus

; APPLICANT: Kroger, Burkhard

; APPLICANT: Schroeder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberhauser, Gregor

; APPLICANT: Haefner, Stefan

; APPLICANT: Klopprogge, Corinna

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING REGULATORY

; TITLE OF INVENTION: PROTEINS INVOLVED IN THE PRODUCTION OF METHIONINE

; FILE REFERENCE: EGI-123CP2

; CURRENT APPLICATION NUMBER: US/10/307,138

; CURRENT FILING DATE: 2002-11-29

; PRIOR APPLICATION NUMBER: 09/602874

; PRIOR FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: 60/422618

; PRIOR FILING DATE: 2002-10-30

; PRIOR APPLICATION NUMBER: 60/141031

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: 60/142690

; PRIOR FILING DATE: 1999-07-01

; PRIOR APPLICATION NUMBER: 60/151251

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19930476.9

; PRIOR FILING DATE: 1999-07-01

; PRIOR APPLICATION NUMBER: DE 19931419.5

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931420.9

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19932122.1

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932128.0

; PRIOR FILING DATE: 1999-07-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 15

; LENGTH: 2772

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: vector

; OTHER INFORMATION: PINTTEGRATIV deitax02910 for knock-out of

; OTHER INFORMATION: positive regulator of methionine biosynthesis

; OTHER INFORMATION: rx02910

; NAME/KEY: misc feature

; LOCATION: (457)..(1248)

; OTHER INFORMATION: kanamycin resistance gene form Tn5

; FEATURE:

NAME/KEY: misc feature
LOCATION: (1515)..(2375)
OTHER INFORMATION: Ori from pMB for replication in E.coli
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(61)
OTHER INFORMATION: fragment of rxa02910 (cont. from bp 2772 in
OTHER INFORMATION: circular plasmid)
FEATURE:
NAME/KEY: misc feature
LOCATION: (2407)..(2772)
OTHER INFORMATION: fragment of rxa02910 (1st part to be con. from bp
OTHER INFORMATION: 1 to 61 in circular plasmid)
US-10-307-138-15

Query Match 100.0%; Score 1001; DB 13; Length 2772;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCCGGGCTCCGGAAACAGATCCGAAGCCCAACCTTTCATAG 60
DB 1390 CGCTGGAGGATCATCCAGCCGGGCTCCGGAAACAGATCCGAAGCCCAACCTTTCATAG 1331

QY 61 AAGCGCGCGTGGAATCGAATCTCGTATGCGAGGTTGGCGTCCGTTGGTGGTCAAT 120
DB 1330 AAGCGCGCGTGGAATCGAATCTCGTATGCGAGGTTGGCGTCCGTTGGTGGTCAAT 1271

QY 121 TCGAACCCCAAGATCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 180
DB 1270 TCGAACCCCAAGATCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 1211

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGGAGCGGTGAGCCATTCGCGCCAA 240
DB 1210 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGGAGCGGTGAGCCATTCGCGCCAA 1151

QY 241 GCTCTTCAGCAATATCAGCGGTAGCAAGCTATGCTGCTGATGCGGTGCGGCTTGGCGTGG 300
DB 1150 GCTCTTCAGCAATATCAGCGGTAGCAAGCTATGCTGCTGATGCGGTGCGGCTTGGCGTGG 1091

QY 301 GCGGCGCAAGTCAGTGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGGCAAGC 360
DB 1090 GCGGCGCAAGTCAGTGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGGCAAGC 1031

QY 361 AGCATCGCATGGGTTCAGACAGATCTCTCGGTCGGGATGCGGCTTGGCGTGGCGTGG 420
DB 1030 AGCATCGCATGGGTTCAGACAGATCTCTCGGTCGGGATGCGGCTTGGCGTGGCGTGG 971

QY 421 CGAACAGTTTCGGTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
DB 970 CGAACAGTTTCGGTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 911

QY 481 GACCGGCTTCATTCGAGTACGTGCTCGATCGATGTTTTCGTTGGTGGTGGTGGATG 540
DB 910 GACCGGCTTCATTCGAGTACGTGCTCGATCGATGTTTTCGTTGGTGGTGGTGGATG 851

QY 541 GCGAGTACCGGATCAAGCGTATCAGCGCGGATTCGATCGATGATGATGATGATGATG 600
DB 850 GCGAGTACCGGATCAAGCGTATCAGCGCGGATTCGATCGATGATGATGATGATGATG 791

QY 601 TCTGCGAGGAGCAAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 660
DB 790 TCTGCGAGGAGCAAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 731

QY 661 GCGAGTTCCTTCGCGCTTCAGTGACAGTTCGATGATGATGATGATGATGATGATGAT 720
DB 730 GCGAGTTCCTTCGCGCTTCAGTGACAGTTCGATGATGATGATGATGATGATGATGAT 671

QY 721 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 670 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611

QY 781 CGGCTTTGACAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

DB 610 CGGCTTTCACAAAAAAGAACCGGCGCGCCCTCGCTGACAGCGGAAACACGCGGCATCAG 551
QY 841 AGCAGCCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
DB 550 AGCAGCCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 491
QY 901 GAGAACCTGCGTGCAATCCATCTTGTTCATATCGGAAACGATCTCTCTGTTCTT 960
DB 490 GAGAACCTGCGTGCAATCCATCTTGTTCATATCGGAAACGATCTCTCTGTTCTT 431
QY 961 GATCAGATCTTGTATCCCTCGCGCATCAGATCTTGGCGG 1001
DB 430 GATCAGATCTTGTATCCCTCGCGCATCAGATCTTGGCGG 390

RESULT 3
US-09-027-287-50
; Sequence 50, Application US/09027287A
; Patent No. US20020064869A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ullrich, Stephen
; TITLE OF INVENTION: Apoptosis Inducing Molecule II
; FILE REFERENCE: 1488 0650004
; CURRENT APPLICATION NUMBER: US/09/027,287A
; CURRENT FILING DATE: 1998-02-20
; EARLIER APPLICATION NUMBER: US 09/003,886
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: US 08/822,953
; EARLIER FILING DATE: 1997-03-21
; EARLIER APPLICATION NUMBER: US 60/030,157
; EARLIER FILING DATE: 1996-10-31
; EARLIER APPLICATION NUMBER: US 60/013,923
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 50
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-09-027-287-50

Query Match 100.0%; Score 1001; DB 9; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGGCTCCCGGAAACGATTCGAAGCCCAACCTTTCATAG 60
DB 2579 CGCTGGAGGATCATCCAGCGGCTCCCGGAAACGATTCGAAGCCCAACCTTTCATAG 2638

QY 61 AAGCGCGCGTGGAATCGAAATCTCTGATGCGAGTGGCGTTCGTTGGTGGTCAAT 120
DB 2639 AAGCGCGCGTGGAATCGAAATCTCTGATGCGAGTGGCGTTCGTTGGTGGTCAAT 2698

QY 121 TCGAACCCCAAGTCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 180
DB 2699 TCGAACCCCAAGTCCCGCTCAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 2758

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGGCGGTGAGCCATTCGCGGCCAA 240
DB 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGGCGGTGAGCCATTCGCGGCCAA 2818

QY 241 GCTCTTCAGCAATATCAGCGGTAGCAAGCTATGCTCTGATAGCGGTGCGGCACACCA 300
DB 2819 GCTCTTCAGCAATATCAGCGGTAGCAAGCTATGCTCTGATAGCGGTGCGGCACACCA 2878

QY 301 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGGCAAGC 360
DB 2879 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGGCAAGC 2938

361 AGGATCCGATCGGTCACGACGAGATCCTCGCGTCGGCATCGCGCTTGAGCTGG 420
Db
2939 AGGATCGCATGGGTACGACGAGATCCTCGCGTCGGCATCGCGCTTGAGCTGG 2998
Qy 421 CGAACAGTTGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
Db 2999 CGAACAGTTGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 3058
Qy 481 GACCGGCTCCATCCGAGTACGTCGCTCGATGAGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTCCATCCGAGTACGTCGCTCGATGAGATGTTTCGCTTGGTGGTGAATG 3118
Qy 541 GGCAGTACGCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATGAT 600
Db 3119 GGCAGTACGCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATGAT 3178
Qy 601 TCTCGCAGCAGCAGTGAATGACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 660
Db 3179 TCTCGCAGCAGCAGTGAATGACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 3238
Qy 661 GGCAGTACGCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATGAT 720
Db 3239 GGCAGTACGCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATGAT 3298
Qy 721 TGCAGCAGCAGTGAATGACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 780
Db 3299 TGCAGCAGCAGTGAATGACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 3358
Qy 781 CGGCTCTTGACAAAGAACCGGCGGCTGAGTACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 840
Db 3359 CGGCTCTTGACAAAGAACCGGCGGCTGAGTACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 3418
Qy 841 AGCAGCGATGTTCTGTTGTCGAGTATGATGCGGATGATGCGGATGATGCGGAT 900
Db 3419 AGCAGCGATGTTCTGTTGTCGAGTATGATGCGGATGATGCGGATGATGCGGAT 3478
Qy 901 GAGAACCTGGTGCATTCATCTGTTGATGATGATGATGATGATGATGATGAT 960
Db 3479 GAGAACCTGGTGCATTCATCTGTTGATGATGATGATGATGATGATGATGAT 3538
Qy 961 GATCAGATCTGATCCCTGCGCCATCAGATCCTTGGCGGC 1001
Db 3539 GATCAGATCTGATCCCTGCGCCATCAGATCCTTGGCGGC 3579

RESULT 4

US-09-252-656B-50
; Sequence 50, Application US/09252656B
; Patent No. US20020081647A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Zhang, Jun
; APPLICANT: Ullrich, Stephen
; APPLICANT: Zhai, Yifan
; TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
; FILE REFERENCE: 1488 0650006
; CURRENT APPLICATION NUMBER: US/09/252,656B
; CURRENT FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 60/075,409
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 09/027,287
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 09/003,886
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: US 08/822,953
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/013,923
; PRIOR FILING DATE: 1996-03-22
; PRIOR APPLICATION NUMBER: US 60/030,157
; PRIOR FILING DATE: 1996-10-31

; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.11
; SEQ ID NO 50
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pHE4-5 expression vector
US-09-252-656B-50

Query Match 100.0%; Score 1001; DB 9; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGTGGAGGATCATCCAGCGGCTCCGGAACCGATTCGGAAGCCAAACCTTTTCATAG 60
Db 2579 CGTGGAGGATCATCCAGCGGCTCCGGAACCGATTCGGAAGCCAAACCTTTTCATAG 2638
Qy 61 AAGCGCGGCTGGAATCGAATCTCGGTGATGGCAGGTTGGCGCTCGTTGGTGGTCAATT 120
Db 2639 AAGCGCGGCTGGAATCGAATCTCGGTGATGGCAGGTTGGCGCTCGTTGGTGGTCAATT 2698
Qy 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758
Qy 181 GCGNATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGGCCAATTCGCCGCCAA 240
Db 2759 GCGNATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGGCCAATTCGCCGCCAA 2818
Qy 241 GCTCTTCAGCAATATCAGGCTGATCCAGCAAGCGGCAATTTCCACCATGATATTCGCAAGC 300
Db 2819 GCTCTTCAGCAATATCAGGCTGATCCAGCAAGCGGCAATTTCCACCATGATATTCGCAAGC 2878
Qy 301 GCGGCGCACAGTGCATGATTCAGAAAGCGGCAATTTCCACCATGATATTCGCAAGC 360
Db 2879 GCGGCGCACAGTGCATGATTCAGAAAGCGGCAATTTCCACCATGATATTCGCAAGC 2938
Qy 361 AGGCATCGCATCGGTGACGAGAGATCCTCGCGTCGGGATGCGGCGCTTGAGCTGG 420
Db 2939 AGGCATCGCATCGGTGACGAGAGATCCTCGCGTCGGGATGCGGCGCTTGAGCTGG 2998
Qy 421 CGAACAGTTGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
Db 2999 CGAACAGTTGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 3058
Qy 481 GACCGGCTCCATCCGAGTACGTCGCTCGATGAGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTCCATCCGAGTACGTCGCTCGATGAGATGTTTCGCTTGGTGGTGAATG 3118
Qy 541 GGCAGTACGCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATGAT 600
Db 3119 GGCAGTACGCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATGAT 3178
Qy 601 TCTCGCAGCAGCAGTGAATGACGAGATTCCTGCGCGGCACTTCGCGCAATAGCA 660
Db 3179 TCTCGCAGCAGCAGTGAATGACGAGATTCCTGCGCGGCACTTCGCGCAATAGCA 3238
Qy 661 GGCAGTACCCCTCCCGCTTCAGTGACAACTGAGCAGATGCTGCGGCAAGCGCGCTCG 720
Db 3239 GGCAGTACCCCTCCCGCTTCAGTGACAACTGAGCAGATGCTGCGGCAAGCGCGCTCG 3298
Qy 721 TGCAGCAGCAGTGAATGACGAGATCCTGCTCGATGATGATGATGATGATGATGAT 780
Db 3299 TGCAGCAGCAGTGAATGACGAGATCCTGCTCGATGATGATGATGATGATGATGAT 3358
Qy 781 CGGCTCTTGACAAAGAACCGGCGGCTGAGTACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 840
Db 3359 CGGCTCTTGACAAAGAACCGGCGGCTGAGTACGAGATCCTGCGCGGCACTTCGCGCAATAGCA 3418
Qy 841 AGCAGCGATGTTCTGTTGTCGAGTATGATGCGGATGATGCGGATGATGCGGAT 900
Db 3419 AGCAGCGATGTTCTGTTGTCGAGTATGATGCGGATGATGCGGATGATGCGGAT 3478

QY 901 GAGACCTCGCTGCAATCCATCTTGTTCATCATCGGAAACGATCCCTCATCTCTCTT 960
DB 3479 GAGACCTCGCTGCAATCCATCTTGTTCATCATCGGAAACGATCCCTCATCTCTCTT 3538
QY 961 GATCAGATCTTGCATCCCTCGGCAATCATGATCCCTTGGCGGC 1001
DB 3539 GATCAGATCTTGCATCCCTCGGCAATCATGATCCCTTGGCGGC 3579

RESULT 5

US-09-972-912-8
; Sequence 8, Application US/09972912
; Patent No. US20020110867A1
; GENERAL INFORMATION:
; APPLICANT: SOPPET, DANIEL R.
; RUBEN, STEVEN M.
; TITLE OF INVENTION: CARDIAC AND PANCREATIC PROTEIN AND GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/972,912
; FILING DATE: 10-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/049,022
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0620001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; LENGTH: 3974 base pairs
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-972-912-8

Query Match 100.0%; Score 1001; DB 10; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCAGCGCGCTCCGGAACGATTCGGAAGCCCAACCTTTCATAG 60
DB 2579 CGCTGGAGGATCATCAGCGCGCTCCGGAACGATTCGGAAGCCCAACCTTTCATAG 2638
QY 61 AAGCGCGCTGGAATCGAATCTCGTATGCGAGGTGGCGTGGCGTGGCGTGGCGTGGCGT 120
DB 2639 AAGCGCGCTGGAATCGAATCTCGTATGCGAGGTGGCGTGGCGTGGCGTGGCGTGGCGT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGGATCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGGATCGCT 2759
QY 181 GCGAATCGGAGGCGGATCCGTAAGCAGAGGAGCGGTACGCCCATTCGCGCCCAA 240

DB 2759 GCGAATCGGAGGCGGATACCGTAAAGCACGAGGAGCGGTACGCCCATTCGCGCCCAA 2818
QY 241 GCTCTTCAGCAATATCACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCGCACACCCA 300
DB 2819 GCTCTTCAGCAATATCACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCGCACACCCA 2878
QY 301 GCGGCGCACAGTGCATGAATCCAGAAAAGCGGCCCAATTTCCACCATGATATTCGGAAGC 360
DB 2879 GCGGCGCACAGTGCATGAATCCAGAAAAGCGGCCCAATTTCCACCATGATATTCGGAAGC 2938
QY 361 AGGCATCGCATGGGTTCAGCAGAGATCTCGCCCTCGGGATCGCGCTTGGCCTTGGCCTGG 420
DB 2939 AGGCATCGCATGGGTTCAGCAGAGATCTCGCCCTCGGGATCGCGCTTGGCCTTGGCCTGG 2998
QY 421 CGAACAGTTCCGCTCGCGCGAGCCCTTGATGTCTTCTTCCAGATCATCTCTGATCGACAA 480
DB 2999 CGAACAGTTCCGCTCGCGCGAGCCCTTGATGTCTTCTTCCAGATCATCTCTGATCGACAA 3058
QY 481 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATCGATGTTTGGTGGTGGTGGTGGTGG 540
DB 3059 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATCGATGTTTGGTGGTGGTGGTGGTGG 3118
QY 541 GCGAGTACCGGATCAAGCGGTATGACGCGCGCGCAATGCAATGCAATGCAATGCAATGCA 600
DB 3119 GCGAGTACCGGATCAAGCGGTATGACGCGCGCGCAATGCAATGCAATGCAATGCAATG 3178
QY 601 TCTCGCAGGAGCAAGGTGAGATGACGAGGATGCTTCCCGCGCACTTTCGCGCCCAATAGCA 660
DB 3179 TCTCGCAGGAGCAAGGTGAGATGACGAGGATGCTTCCCGCGCACTTTCGCGCCCAATAGCA 3238
QY 661 GCGAGTCCCTTCCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 720
DB 3239 GCGAGTCCCTTCCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 3298
QY 721 TGGCAGCCACGATAGCGCGCTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 780
DB 3299 TGGCAGCCACGATAGCGCGCTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 3358
QY 781 CGGTCTTGACAAAAGAACCGCGCGCGCTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 840
DB 3359 CGGTCTTGACAAAAGAACCGCGCGCGCTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 3418
QY 841 AGCAGCCGATTTGCTTGTGTCGCAATAGCGCAATAGCGCTTCCACCCCAAGCGCGCG 900
DB 3419 AGCAGCCGATTTGCTTGTGTCGCAATAGCGCAATAGCGCTTCCACCCCAAGCGCGCG 3478
QY 901 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAACGATCCCTCATCTCTCTCTT 960
DB 3479 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAACGATCCCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGCATCCCTCGGCGCATCAGATCCCTTGGCGGC 1001
DB 3539 GATCAGATCTTGCATCCCTCGGCGCATCAGATCCCTTGGCGGC 3579

RESULT 6

US-09-902-684-14
; Sequence 14, Application US/09902684
; Patent No. US20020127640A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR INHIBITOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Qy		601	TCTCGCAGGAGCAAGTGAGATGCACAGAGATCTTGGCCCCGGCACTTCGCCCAATAGCA	660
Dδ		3179	TCTCGCAGGAGCAAGTGAGATGCACAGAGATCTTGGCCCCGGCACTTCGCCCAATAGCA	3238

QY 661 GCCAGTCCCTTCGCCGTTTCA GTGA CAACGTCTCGAGCA CAGCTGC GC GAAGNACGCCCGTCG 720
| | | | |
Db 3239 GCCAGTCCCTTCGCCGTTTCA GTGA CAACGTCTCGAGCA CAGCTGC GC GAAGNACGCCCGTCG 3298

Qy	721	TGSCCAGCCACGATAGCCGGCTCGTCTCTGCGATTTCATTCAGGGCACCGACAGGT	780
Db	3299	TGSCCAGCCACGATAGCCGGCTCGTCTCTGCGATTTCATTCAGGGCACCGACAGGT	3358

Qy 781 CGGCTTTGACAAAAGAACCGGGGGCCCCCTGCGCTGACAGCGCGGAACACGCGCGGCATCAG 840

Db 3359 CGGCTTTGACAAAAGAACCGGGGGCCCCCTGCGCTGACAGCGCGGAACACGCGCGGCATCAG 3418

QY 841 AGCAGCGGATTCGTCTGTTTGCCCAAGTACGCCGAATAGCCTTCCACCCCAAGCGGC 900
| | | | |
Db 3419 AGCAGCGGATTCGTCTGTTTGCCCAAGTACGCCGAATAGCCTTCCACCCCAAGCGGC 3478

Qy 901 GAGAACTGCGGCAATCCATCTTTTCAATCATGCGAAAAGATCCTCATCTGTCTCTTT 960
Dβ 3479 GAGAACTGCGGCAATCCATCTTTTCAATCATGCGAAAAGATCCTCATCTGTCTCTTT 3538	

Qy 961 GATCAGATCTTGATCCCTGGGCCATCAGATCCTTGGCGGC 1001
 ||| ||| ||| ||| |||
Db 3539 GATCAGATCTTGATCCCTGGGCCATCAGATCCTTGGCGGC 3579

RESULT 7
US-09-935-726-16
Sentenced 16

```

; Sequence 1b, Application US/09935726
; Publication No. US2003008357A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Hu, Jin-Shan
; ADDRESS: Taipei, Taiwan, R.O.C.
;

```

```

: AFFILIANT: Craig, Roben
:
: APPLICANT: Cao, Liang
:
: TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
:
: FILE REFERENCE: PF1123D1C1
:
: CURRENT APPLICATION NUMBER: 10/000000

```

7 CURRENT APPLICATION NUMBER: US/09/935,726
 7
 7 CURRENT FILING DATE: 2001-08-24
 7
 7 PRIOR APPLICATION NUMBER: 09/438,538
 7
 7 PRIOR FILING DATE: 1999-11-12
 7
 7 NUMBER OF SEQ TO VOC: 25
 7

```

; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3974
; TYPE: DNA

```

```

/ life: DNA
/ ORGANISM: Escherichia coli
US-09-935-726-16
Query Match 100.0% 100.0% 100.0% 100.0% 100.0%

```

Query Match 100.0%; Score 1001; DB 11; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

61 AAGCGCGGTGGAAACGAAATCTCGTATGCCAGTTGGCGTCGCTTGCGTCATT 120
2639 AAGCGCGGTGGAAACGAAATCTCGTATGCCAGTTGGCGTCGCTTGCGTCATT 2698
121 TCGAACCCCAAGACGACCGCGCCTCATCACACGCGGCATCACACGCGCGCGCGCGT 120

[illegible]

181 GCGAAATGGGAGCGCGGATACCGTAAAGACAGAGGAAGCGGTAGCCCATTCGCGCGCAA 240
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992

5' 3' 300

Db 2819 GCTCTTACGAATATACCGGTTAGCAAGCGTATGTCTCTGATAGCGTCCGCCACACCA 2878
QY 301 GCCGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCCGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGGGTGTCACGACGATCCTCGCGTGGGATGCGCGCTTGAGCTGG 420
Db 2939 AGGCATCGCCATGGGTGTCACGACGATCCTCGCGTGGGATGCGCGCTTGAGCTGG 2998
QY 421 CGAACAGTTCCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGCTTCCATCCGATGATGCTGCTCGATGCGATGTTTCGCTGGTGGTGGTGG 540
Db 3059 GACCGCTTCCATCCGATGATGCTGCTCGATGCGATGTTTCGCTGGTGGTGGTGG 3118
QY 541 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 600
Db 3119 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3178
QY 601 TCTGGCAGGACGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 3179 TCTGGCAGGACGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3238
QY 661 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 720
Db 3239 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3298
QY 721 TGCCGAGCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 3299 TGCCGAGCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
QY 781 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 3359 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
QY 841 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
QY 901 GAGAACCTGCTGCAATCCATCTGTTCAATCATGCGAAGCATGCTGCTGCTGCTGCTGCT 960
Db 3479 GAGAACCTGCTGCAATCCATCTGTTTCAATCATGCGAAGCATGCTGCTGCTGCTGCT 3538
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCCATCAGATCCTTTGGCGGC 3579

RESULT 8
US-09-105-470-25
; Sequence 25, Application US/09105470
; Publication No. US20030022275A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Kenneth C.
; APPLICANT: He, Wei-Wu
; TITLE OF INVENTION: Human NK-3 Related Prostate Specific
; TITLE OF INVENTION: Gene-1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,470
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/051,080
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0790001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; US-09-105-470-25

Query Match 100.0%; Score 1001; DB 11; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60
Db 2579 CCGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGCGGTGGAATCGAATCTCGTATGCGAGTTGGGCGTCTGCTGCTGCTGCTGCT 120
Db 2639 AAGCGCGCGGTGGAATCGAATCTCGTATGCGAGTTGGGCGTCTGCTGCTGCTGCTGCT 2698
QY 121 TCGAACCCCGAGTCCCCCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGGATCGCCT 180
Db 2699 TCGAACCCCGAGTCCCCCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGGATCGCCT 2758
QY 181 GCGAATCGGAGCGCGGATAGCGTAAGACGAGAGAGCGGTGAGCCCATTTGCGCGCCAA 240
Db 2759 GCGAATCGGAGCGCGGATAGCGTAAGACGAGAGAGCGGTGAGCCCATTTGCGCGCCAA 2818
QY 241 GCTCTTACGAATATCACGGGTAGCCACGCTATGCTCTGATAGCGTCCGCCACACCCA 300
Db 2819 GCTCTTACGAATATCACGGGTAGCCACGCTATGCTCTGATAGCGTCCGCCACACCCA 2878
QY 301 GCCGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCCGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGGGTGTCACGACGATCCTCGCGTGGGATGCGCGCTTGAGCTGG 420
Db 2939 AGGCATCGCCATGGGTGTCACGACGATCCTCGCGTGGGATGCGCGCTTGAGCTGG 2998
QY 421 CGAACAGTTCCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGCTTCCATCCGATGATGCTGCTCGATGCGATGTTTCGCTGGTGGTGGTGG 540
Db 3059 GACCGCTTCCATCCGATGATGCTGCTCGATGCGATGTTTCGCTGGTGGTGGTGG 3118
QY 541 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 600
Db 3119 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3178
QY 601 TCTGGCAGGACGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 3179 TCTGGCAGGACGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3238
QY 661 GCGAGTTCGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 720

781	QY	CGGTCTTGACAAAAGAAACCGGGGCCCTTCGCTGACAGCCGGAAACGCGCGGCATCAG	840
3359	Db	CGGTCTTGACAAAAGAAACCGGGGCCCTTCGCTGACAGCCGGAAACGCGCGGCATCAG	3418
841	QY	AGCAGCCGATGTCTGTGTGTGCCAGTCATAGCCGAATAGCTCTTCCACCCAAAGGGCGC	900
3419	Db	AGCAGCCGATGTCTGTGTGTGCCAGTCATAGCCGAATAGCTCTTCCACCCAAAGGGCGC	3478
901	QY	GAGAACCTCGCGTCAATCTTGTTTCAATCATGCGAAACGATCCTCATCTGTCTCTT	960
3479	Db	GAGAACCTCGCGTCAATCTTGTTTCAATCATGCGAAACGATCCTCATCTGTCTCTT	3538
961	QY	GATCAGATCTTGATCCCTCGGCCCATCAGATCTTGGCGC	1001
3539	Db	GATCAGATCTTGATCCCTCGGCCCATCAGATCTTGGCGC	3579

RESULT 13

RESULI 13
US-10-614-275-25
Sequence 25, Application US/10614275
Publication No. US20040014668A1
GENERAL INFORMATION:
APPLICANT: Carter, Kenneth C.
He, Wei-Wu
TITLE OF INVENTION: Human NK-3 Related Prostate Specific
Gene-1
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA

```

/ COMPUTER READABLE FORM:
/ ZIP: 200005
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/VMS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/

```

?
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/10/614,275
? FILING DATE: 08-Jul-2003
? CLASSIFICATION: <Unknown>
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/09/105,470B
? FILING DATE: 26-Jun-1998
? APPLICATION NUMBER: US 60/051,080
? FILING DATE: 27-JUN-1997
? ATTORNEY/AGENT INFORMATION:
?

ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0790001
TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 3974 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-614-275-25

QY	61	AAGCGCGCGTGGAAATCGAAATCTCGTGTATGGCAGGTTGGCGTCGCTTGGTCGGTCAATT	120
Db	2639	AAGCGCGCGTGGAAATCGAAATCTCGTGTATGGCAGGTTGGCGTCGCTTGGTCGGTCAATT	2698
QY	121	TCGAACCCACAGAGTCCGCTCAGAAGAACTCGTCACAGAGCGCATAGAAGCGCATGCGCT	180
Db	2699	TCGAACCCACAGAGTCCGCTCAGAAGAACTCGTCACAGAGCGCATAGAAGCGCATGCGCT	2758
QY	181	GCGAATCGGAGCGGGGATACCGTAAAGCAACGAGGAAGCGGTTCAGCCCATTCGCCGCGCAA	240
Db	2759	GCGAATCGGAGCGGGGATACCGTAAAGCAACGAGGAAGCGGTTCAGCCCATTCGCCGCGCAA	2818
QY	241	GCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGTCCTGATAGCGGTCCGCACACCCCA	300
Db	2819	GCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGTCCTGATAGCGGTCCGCACACCCCA	2878
QY	301	GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC	360
Db	2879	GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC	2938
QY	361	AGCAGTCGCATGGGTACAGACGAGATCTTCGCGGTTCGGCAGTCGCGCTTCAGGCTGG	420
Db	2939	AGCAGTCGCATGGGTACAGACGAGATCTTCGCGGTTCGGCAGTCGCGCTTCAGGCTGG	2998
QY	421	CGAACCATGTCGGGTGGCGAGCGCCCTGATGCTCTTTCGTCACAGATCATCTGTATGCACAA	480
Db	2999	CGAACCATGTCGGGTGGCGAGCGCCCTGATGCTCTTTCGTCACAGATCATCTGTATGCACAA	3058
QY	481	GACCGGTTCCATCCGAGTACGTCTCGTCGATCGATGTTTCGCTTGGTGGTTCGAATG	540
Db	3059	GACCGGTTCCATCCGAGTACGTCTCGTCGATCGATGTTTCGCTTGGTGGTTCGAATG	3118
QY	541	GGCAGGTAGCCGGATCAAGCGTATGCAGCGCGCGCATTTGCATCAGCCATGATGGATACCT	600
Db	3119	GGCAGGTAGCCGGATCAAGCGTATGCAGCGCGCGCATTTGCATCAGCCATGATGGATACCT	3178
QY	601	TCTCGGACGAGCAAGGTGAGATGACAGGAGATCTCTGCCCCGCGCATCTGCCCAATAGCA	660
Db	3179	TCTCGGACGAGCAAGGTGAGATGACAGGAGATCTCTGCCCCGCGCATCTGCCCAATAGCA	3238
QY	661	GCCAGTCCCTTCCCGCTTCAGTCAACACGTTCGAGCACAGTCGCGCAAGGAACGCCCGCTCG	720
Db	3239	GCCAGTCCCTTCCCGCTTCAGTCAACACGTTCGAGCACAGTCGCGCAAGGAACGCCCGCTCG	3298
QY	721	TGGCCACGACGATAGCGCGCTCGCTCTCTGCGATGCTATTCAGGGCACCGGACAGGT	780
Db	3299	TGGCCACGACGATAGCGCGCTCGCTCTCTGCGATGCTATTCAGGGCACCGGACAGGT	3358
QY	781	CGGTCTTGACAAAAGAACCGGGCGCCCTTGCGCTCAGACGCGGAACACGGCGGCATCAG	840
Db	3359	CGGTCTTGACAAAAGAACCGGGCGCCCTTGCGCTCAGACGCGGAACACGGCGGCATCAG	3418
QY	841	AGCAGCGGATGCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCGAGCGGCGG	900
Db	3419	AGCAGCGGATGCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCGAGCGGCGG	3478
QY	901	GAGAACTCGTGCATCCATCTGTTCATCATGCGAAACGATCCTCATCTCTGTCTCTT	960
Db	3479	GAGAACTCGTGCATCCATCTGTTCATCATGCGAAACGATCCTCATCTCTGTCTCTT	3538
QY	961	GATCAGATCTTGATCCCTCGCGCATCAGATCCTTGGCGGC	1001
Db	3539	GATCAGATCTTGATCCCTCGCGCATCAGATCCTTGGCGGC	3579

RESULT 14
US-10-260-270-4
; Sequence 4, Application US/10260270
; Publication NO. US20030143192A1
; GENERAL INFORMATION:
; APPLICANT: Salcedo, Theodora W.
; APPLICANT: Patel, Vikram, P.

APPLICANT: Nibbs, Robert J.B.
APPLICANT: Graham, Gerard J.
TITLE OF INVENTION: CHEMOKINE BETA-7
FILE REFERENCE: 1488.033000F
CURRENT APPLICATION NUMBER: US/10/260,270
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: US/09/437,602
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: US 60/107,801
PRIOR FILING DATE: 1998-11-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent in version 3.0
SEQ ID NO 4
LENGTH: 3974
TYPE: DNA
ORGANISM: Homo sapiens
US-10-260-270-4

Query Match 100.0%; Score 1001; DB 13; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 2638

QY 61 AAGCGCGGTGGAATCGAATCTCGTATCGATGCGGTTGGCGCTGCTTGGTCTGCTATT 120
Db 2639 AAGCGCGGTGGAATCGAATCTCGTATCGATGCGGTTGGCGCTGCTTGGTCTGCTATT 2698

QY 121 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 180
Db 2699 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 2758

QY 181 GCGATCGGCGCGGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 240
Db 2759 GCGATCGGCGCGGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 2818

QY 241 GCTCTTACGAATATCAGGCGTAGCCAGCGTATGCTGATAGCGTCCGCGCACCCCA 300
Db 2819 GCTCTTACGAATATCAGGCGTAGCCAGCGTATGCTGATAGCGTCCGCGCACCCCA 2878

QY 301 GCGGCGCAGTCGATGATGATCAGAGAGCGGCGATTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCGGCGCAGTCGATGATGATCAGAGAGCGGCGATTTCCACCATGATATTCGGCAAGC 2938

QY 361 AGGCATCGCATGGTGCAGCAGAGATCTCGCGCTGGCGATCGCGCTTGGAGCTGG 420
Db 2939 AGGCATCGCATGGTGCAGCAGAGATCTCGCGCTGGCGATCGCGCTTGGAGCTGG 2998

QY 421 CGACAGTTCGGCTGGCGGAGCCCGTATGCTCTTCTCGATCATCTGATCGACAA 480
Db 2999 CGACAGTTCGGCTGGCGGAGCCCGTATGCTCTTCTCGATCATCTGATCGACAA 3058

QY 481 GACCGGCTTCATCGAGTACGTCTCGCTGATGCGATGCTTTCGCTGGTGGTCTGATG 540
Db 3059 GACCGGCTTCATCGAGTACGTCTCGCTGATGCGATGCTTTCGCTGGTGGTCTGATG 3118

QY 541 GCGAGTAGCGGATCAAGCGTATGAGCCCGCATTTGATCAGCGATGATGATGATCTT 600
Db 3119 GCGAGTAGCGGATCAAGCGTATGAGCCCGCATTTGATCAGCGATGATGATGATCTT 3178

QY 601 TCTCGGAGGAGCAGGTGATGATGAGAGATCTCGCGCGGACATTCGCCCAATAGCA 660
Db 3179 TCTCGGAGGAGCAGGTGATGATGAGAGATCTCGCGCGGACATTCGCCCAATAGCA 3238

QY 661 GCCAGTCCCTTCCCGCTTCACTGACAGCTCGACACAGTTCGCGAGGAGCCCGCTG 720
Db 3239 GCCAGTCCCTTCCCGCTTCACTGACAGCTCGACACAGTTCGCGAGGAGCCCGCTG 3298

QY 721 TGGCCAGCAGATAGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 3299 TGGCCAGCAGATAGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358

QY 781 CGCTCTTGACAAAAGAACCGGCGCGCTTGCTGCTGACAGCGGACACAGCGGCGATCAG 840
Db 3359 CGCTCTTGACAAAAGAACCGGCGCGCTTGCTGCTGACAGCGGACACAGCGGCGATCAG 3418

QY 841 AGCAGCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGCAGCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478

QY 901 GAGAAGCTCGTGCATTCATCTGTTCAATCATGCGAAGAGATCTCTCTCTCTCTCTCT 960
Db 3479 GAGAAGCTCGTGCATTCATCTGTTCAATCATGCGAAGAGATCTCTCTCTCTCTCTCT 3538

QY 961 GATCAGATCTTGTATCCCTGCGCGCATCAGATCTTGGCGGC 1001
Db 3539 GATCAGATCTTGTATCCCTGCGCGCATCAGATCTTGGCGGC 3579

RESULT 15
US-10-197-844-9
Sequence 9, Application US/10197844
Publication No. US2003016689A1
GENERAL INFORMATION:
APPLICANT: CHOPRA, ARVIND
OLSEN, HENRIK
GENIZ, STEVEN M.
RUBIN, STEVEN M.
TITLE OF INVENTION: MYELIN OLIGODENDROCYTE GLYCOPROTEIN-LIKE
PROTEIN (MOGP) AND METHODS OF USE
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/197,844
FILING DATE: 19-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/015,963
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/035,445
FILING DATE: 30-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0520001/EKS/RCM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3974 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-197-844-9

Query Match 100.0%; Score 1001; DB 13; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 2638

QY 61 AAGCGCGGTGGAATCGAATCTCGTATCGATGCGGTTGGCGCTGCTTGGTCTGCTATT 120
Db 2639 AAGCGCGGTGGAATCGAATCTCGTATCGATGCGGTTGGCGCTGCTTGGTCTGCTATT 2698

QY 121 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 180
Db 2699 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 2758

QY 181 GCGATCGGCGCGGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 240
Db 2759 GCGATCGGCGCGGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 2818

QY 241 GCTCTTACGAATATCAGGCGTAGCCAGCGTATGCTGATAGCGTCCGCGCACCCCA 300
Db 2819 GCTCTTACGAATATCAGGCGTAGCCAGCGTATGCTGATAGCGTCCGCGCACCCCA 2878

QY 301 GCGGCGCAGTCGATGATGATCAGAGAGCGGCGATTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCGGCGCAGTCGATGATGATCAGAGAGCGGCGATTTCCACCATGATATTCGGCAAGC 2938

QY 361 AGGCATCGCATGGTGCAGCAGAGATCTCGCGCTGGCGATCGCGCTTGGAGCTGG 420
Db 2939 AGGCATCGCATGGTGCAGCAGAGATCTCGCGCTGGCGATCGCGCTTGGAGCTGG 2998

QY 421 CGACAGTTCGGCTGGCGGAGCCCGTATGCTCTTCTCGATCATCTGATCGACAA 480
Db 2999 CGACAGTTCGGCTGGCGGAGCCCGTATGCTCTTCTCGATCATCTGATCGACAA 3058

QY 481 GACCGGCTTCATCGAGTACGTCTCGCTGATGCGATGCTTTCGCTGGTGGTCTGATG 540
Db 3059 GACCGGCTTCATCGAGTACGTCTCGCTGATGCGATGCTTTCGCTGGTGGTCTGATG 3118

QY 541 GCGAGTAGCGGATCAAGCGTATGAGCCCGCATTTGATCAGCGATGATGATGATCTT 600
Db 3119 GCGAGTAGCGGATCAAGCGTATGAGCCCGCATTTGATCAGCGATGATGATGATCTT 3178

QY 601 TCTCGGAGGAGCAGGTGATGATGAGAGATCTCGCGCGGACATTCGCCCAATAGCA 660
Db 3179 TCTCGGAGGAGCAGGTGATGATGAGAGATCTCGCGCGGACATTCGCCCAATAGCA 3238

QY 661 GCCAGTCCCTTCCCGCTTCACTGACAGCTCGACACAGTTCGCGAGGAGCCCGCTG 720
Db 3239 GCCAGTCCCTTCCCGCTTCACTGACAGCTCGACACAGTTCGCGAGGAGCCCGCTG 3298

QY 721 TGGCCAGCAGATAGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 3299 TGGCCAGCAGATAGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358


```
Db 2579 CGTGAGGATCATCCAGCGCGTCCCGAAGAGATTCGAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGCGTGGATCGAAATCTCGTATGCGAGTTGGCGTGGCTTGGTTCGCTCAT 120
Db 2639 AAGCGCGCGTGGATCGAAATCTCGTATGCGAGTTGGCGTGGCTTGGTTCGCTCAT 2698
QY 121 TCGAAACCCAGAGTCCCGCTCAGAAGACTCGTCAAGAGGCGATAGAGGCGATGGCT 180
Db 2699 TCGAAACCCAGAGTCCCGCTCAGAAGACTCGTCAAGAGGCGATAGAGGCGATGGCT 2758
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGACAGAGAGCGGTCAAGCCCATTCGCGCCAA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGACAGAGAGCGGTCAAGCCCATTCGCGCCAA 2818
QY 241 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCA 300
Db 2819 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCA 2878
QY 301 GCCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGGGTACAGAGAGATCCTCGCGTCGGGATGCGCGCTTGAGCCTGG 420
Db 2939 AGGCATCGCCATGGGTACAGAGAGATCCTCGCGTCGGGATGCGCGCTTGAGCCTGG 2998
QY 421 CGAACAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCATTCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 540
Db 3059 GACCGGCTTCATTCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 3118
QY 541 GCGAGTAGCGCGATCAAGCGTATGACGCGCGCCATTTGATCAGCCATGATGATACCTT 600
Db 3119 GCGAGTAGCGCGATCAAGCGTATGACGCGCGCCATTTGATCAGCCATGATGATACCTT 3178
QY 601 TCTCGCAGGAGCAAGGTGAGATCAAGAGATCCTGCCCGGACCTTCGCCCAATAGCA 660
Db 3179 TCTCGCAGGAGCAAGGTGAGATCAAGAGATCCTGCCCGGACCTTCGCCCAATAGCA 3238
QY 661 GCCAGTCCCTTCGCGCTTCAGTGACAGTGCAGCACAGTGGCGAAGGAACGCCCGTCG 720
Db 3239 GCCAGTCCCTTCGCGCTTCAGTGACAGTGCAGCACAGTGGCGAAGGAACGCCCGTCG 3298
QY 721 TGGCCAGCCAGATAGCGCGCTGCTCTGCTCGAGTTCATTTCAGGGCACCGGACAGGT 780
Db 3299 TGGCCAGCCAGATAGCGCGCTGCTCTGCTCGAGTTCATTTCAGGGCACCGGACAGGT 3358
QY 781 CGGTCTTGACAAAGAACCGGGCGCCCTGCGCTGACAGCGGAAACACGGCGGCATCAG 840
Db 3359 CGGTCTTGACAAAGAACCGGGCGCCCTGCGCTGACAGCGGAAACACGGCGGCATCAG 3418
QY 841 AGCAGCGGATGTCTGTGTGTCGCCAGTATAGCGGATAGCTCTCCACCCAGCGGCCG 900
Db 3419 AGCAGCGGATGTCTGTGTGTCGCCAGTATAGCGGATAGCTCTCCACCCAGCGGCCG 3478
QY 901 GAGAACCTGCTGCAATCCATCTTGTTCATATGCGAAACGATTCCTCATCTCTCTT 960
Db 3479 GAGAACCTGCTGCAATCCATCTTGTTCATATGCGAAACGATTCCTCATCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCCTTGGCGG 1001
Db 3539 GATCAGATCTTGATCCCTCGCCATCAGATCCCTTGGCGG 3579
```

Search completed: February 15, 2004, 22:26:53
Job time : 425.484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 21:50:21 ; Search time 2636.53 Seconds
(without alignments)
9227.587 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 cctgtgagatccagcc.....gcatcgatccttggcgc 1001

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gas_hum:*
18: em_gas_inv:*
19: em_gas_pln:*
20: em_gas_vrt:*
21: em_gas_fun:*
22: em_gas_mam:*
23: em_gas_mus:*
24: em_gas_pro:*
25: em_gas_rtd:*
26: em_gas_phg:*
27: em_gas_vrl:*
28: gb_gasl:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568	56.7	784	28	AQ876152
2	568	56.7	850	28	AQ875052
3	520	51.9	798	28	AQ876139
4	508	50.7	509	28	AQ447775

5	508	50.7	790	28	AQ876067
6	507	50.6	789	28	AQ361914
7	499	49.9	499	28	AQ398880
8	499	49.9	592	28	AQ449182
9	493	49.3	637	28	AQ447230
10	493	49.3	861	28	AQ875040
11	489	48.9	799	28	AQ876220
12	481	48.1	801	28	AQ875907
13	479	47.9	791	28	AQ876024
14	468	46.8	561	28	AQ447874
15	468	46.8	620	28	AQ398387
16	468	46.8	789	28	AQ876017
17	459	45.9	779	28	AQ875906
18	459	45.9	791	28	AQ876019
19	459	45.9	857	28	AQ875013
20	457	45.7	791	28	AQ876015
21	454	45.4	498	28	AQ398730
22	445	44.5	635	28	AQ447643
23	445	44.5	656	28	AQ447140
24	438	43.8	797	28	AQ876297
25	435	43.5	791	28	AQ876201
26	429	42.9	473	28	AQ397768
27	428	42.8	798	28	AQ875840
28	410	41.0	795	28	AQ876130
29	398	39.8	790	28	AQ875904
30	397	39.7	860	28	AQ875029
31	390	39.0	762	28	AQ873181
32	387	38.7	815	28	AQ876134
33	382	38.2	748	28	AQ501530
34	381	38.1	382	28	AQ448898
35	377	37.7	540	28	AQ874387
36	371	37.1	422	28	BH912281
37	366	36.6	642	28	AQ874213
38	366	36.6	801	28	AQ876288
39	363	36.3	366	28	AQ399385
40	357	35.7	704	28	AQ874178
41	355	35.5	804	28	AQ361943
42	353	35.3	500	12	BM056700
43	351	35.1	734	28	AQ873679
44	349	34.9	779	28	AQ876036
45	348	34.8	705	28	AQ500916

ALIGNMENTS

RESULT 1
AQ876152
LOCUS
DEFINITION
V133H12 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION
AQ876152
VERSION
AQ876152.1 GI:6288396
KEYWORDS
GSS.
SOURCE
Saccharomyces cerevisiae (baker's yeast)
ORGANISM
Saccharomyces cerevisiae
REFERENCE
1 (bases 1 to 784)
AUTHORS
Roos-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deStages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
JOURNAL
Unpublished
COMMENT
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu

QY 421 CGAACAGTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 480
 Db 599 CGAACAGTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 658
 QY 481 GACCGGCTTCATCCGAGTACGTGCTCGTCGATCGGATGTTTCGCTTGGTGGTGAATG 540
 Db 659 GACCGGCTTCATCCGAGTACGTGCTCGTCGATCGGATGTTTCGCTTGGTGGTGAATG 718
 QY 541 GCAGGTAGCCGATCAGCGTATGCG 568
 Db 719 GCAGGTAGCCGATCAGCGTATGCG 746

RESULT 3
 LOCUS A0876139 798 bp DNA linear GSS 08-NOV-1999
 DEFINITION V133G10 mTn-3XHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.

ACCESSION A0876139
 VERSION A0876139.1 GI:6288383
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 798)
 AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A., desStages, S. A., Cheung, K. H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roder, G. S. and Snyder, M.
 TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

JOURNAL Unpublished
 COMMENT Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mTn-3XHA/lacZ insertion.
 Seq primer: GGCCTCTTCTTGTGGAAGTAC
 Class: transposon-tagged

FEATURES
 source
 1..798
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /strain="Y2278 - S288C background, cir(0) rho(0)"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3XHA/lacZ Insertion Library, strain Y2278"
 /note="Vector: pHS56-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3XHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 176 a 230 c 226 g 164 t 2 others
 ORIGIN
 Query Match 51.9%; Score 520; DB 28; Length 798;
 Best Local Similarity 99.8%; Pred No. 9.3e-262;
 Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCTGGAGATCATCCAGCGCGTCCCGAAACGATTCGAAAGCCCACTTCATAG 60
 Db 194 CGCTGGAGATCATCCAGCGCGTCCCGAAACGATTCGAAAGCCCACTTCATAG 253
 QY 61 AAGGCGGCGTGGATCGAATCTCGTATGCGAGTTGGCGTGGTGGTGGTGGTATT 120
 Db 254 AAGGCGGCGTGGATCGAATCTCGTATGCGAGTTGGCGTGGTGGTGGTGGTATT 313

QY 121 TCGAACCCCAAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGCATCGCT 180
 Db 314 TCGAACCCCAAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGCATCGCT 373
 QY 181 GCGAATCGGAGCGGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCCAA 240
 Db 374 GCGAATCGGAGCGGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCCAA 433
 QY 241 GCTCTTACGCAATATCACGGGTAGCCACGCTATGCTCTGATAGCGGTCCGCACACCCA 300
 Db 434 GCTCTTACGCAATATCACGGGTAGCCACGCTATGCTCTGATAGCGGTCCGCACACCCA 493
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGCAAGC 360
 Db 494 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGCAAGC 553
 QY 361 AGGCATCGGCATCGGTACAGACGAGATCCTCGCGGTGCGGCATCGCGGCTTGAGCCTGG 420
 Db 554 AGGCATCGGCATCGGTACAGACGAGATCCTCGCGGTGCGGCATCGCGGCTTGAGCCTGG 613
 QY 421 CGAACAGTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 480
 Db 614 CGAACAGTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 673
 QY 481 GACCGGCTTCATCCGAGTACGTGCTCGATCGGATGTTTCGCTTGGTGGTGAATG 540
 Db 674 GACCGGCTTCATCCGAGTACGTGCTCGATCGGATGTTTCGCTTGGTGGTGAATG 733
 QY 541 GCGAGTAGCCGATCAAGCGTATGCGCG 571
 Db 734 GCGAGTAGCCGATCAAGCGTATGCGCG 764

RESULT 4
 LOCUS A0447775/c 509 bp DNA linear GSS 08-APR-1999
 DEFINITION mgx0011E13f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgx0011E13f, genomic survey sequence.

ACCESSION A0447775
 VERSION A0447775.1 GI:4576912
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea

REFERENCE 1 (bases 1 to 509)
 AUTHORS Yu, Y., Zhu, H., Boyd, C. A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R. A. and Dean, R. A.
 TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCATATAGG
 Class: BAC ends
 High quality sequence stop: 342.
 Location/Qualifiers
 1..509
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgx0011E13f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases

FEATURES
 source

of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

BASE COUNT 97 a 143 c 159 g 109 t 1 others

Query Match 50.7%; Score 508; DB 28; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.8e-255;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 ATCCGGAGCGGATACCGTAAAGACAGAGAGCGGTGAGCCATTCGCGCCAGCTC 244
DB |||||
QY 508 ATCCGGAGCGGATACCGTAAAGACAGAGAGCGGTGAGCCATTCGCGCCAGCTC 449
DB |||||
QY 245 TTCAGCAATATCACGGGTAGCCACGCTATGCTGATAGCGGTCCGCGCACACCCAGCG 304
DB |||||
QY 448 TTCAGCAATATCACGGGTAGCCACGCTATGCTGATAGCGGTCCGCGCACACCCAGCG 389
DB |||||
QY 305 GCCACAGTCGATGATCCAGAAAGCGGCCATTTCCACCATGATATTCGCAAGCAGGC 364
DB |||||
QY 388 GCCACAGTCGATGATCCAGAAAGCGGCCATTTCCACCATGATATTCGCAAGCAGGC 329
DB |||||
QY 365 ATCCGCAATGATCACGAGAGATCTCGCGGTGAGCCATTCGCGCCAGCTC 424
DB |||||
QY 328 ATCCGCAATGATCACGAGAGATCTCGCGGTGAGCCATTCGCGCCAGCTC 269
DB |||||
QY 425 CAGTTCGGTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATCAGCAGACC 484
DB |||||
QY 268 CAGTTCGGTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATCAGCAGACC 209
DB |||||
QY 485 GGCCTTCATCCAGTAGTGTCTGCTGATGATGATTTTCGTTGGTGGTGAATGGCA 544
DB |||||
QY 208 GGCCTTCATCCAGTAGTGTCTGCTGATGATGATTTTCGTTGGTGGTGAATGGCA 149
DB |||||
QY 545 GTAGCCGATCAAGCGTATGACGCGCGCGCATTCGATCAGCAGCATGATGATTTCTC 604
DB |||||
QY 148 GTAGCCGATCAAGCGTATGACGCGCGCGCATTCGATCAGCAGCATGATGATTTCTC 89
DB |||||
QY 605 GCAGAGCAAGTGTGATGACAGGATCTTCGCGCGCATTCGCGCGCATTCGCGCGCAT 664
DB |||||
QY 88 GCAGAGCAAGTGTGATGACAGGATCTTCGCGCGCATTCGCGCGCATTCGCGCGCAT 29
DB |||||
QY 665 GTCCCTTCGCGCTTCAGTGACACGTCG 692
DB |||||
QY 28 GTCCCTTCGCGCTTCAGTGACACGTCG 1

RESULT 5
AQ876067
LOCUS
DEFINITION V132H11 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION AQ876067
VERSION AQ876067.1 GI:6288311
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
REFERENCE Ros-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., desRages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
AUTHORS Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
TITLE Unpublished
JOURNAL

COMMENT

Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTTCCTTCCTTGGAGTAC
Class: transposon-tagged.

FEATURES

source
Location/Qualifiers
1..790

/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/laE_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS96-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS96-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 170 a 225 c 201 g 194 t
ORIGIN

Query Match 50.7%; Score 508; DB 28; Length 790;
Best Local Similarity 100.0%; Pred. No. 1.9e-255;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATTCAGCGCGGTCCGGAAACGATTCGGAAGCCCACTTCATAG 60
DB |||||
QY 282 CGCTGGAGGATCATTCAGCGCGGTCCGGAAACGATTCGGAAGCCCACTTCATAG 341
DB |||||
QY 61 AAGCGCGCGGTGGAATCGAAATCTCGTGTGGCAGGTTCGGCGTTCGTTGGTGGTCA 120
DB |||||
QY 342 AAGCGCGCGGTGGAATCGAAATCTCGTGTGGCAGGTTCGGCGTTCGTTGGTGGTCA 401
DB |||||
QY 121 TCGAACCCAGATCCCGCTCAGAAGAACTCGTCAAGAGGGCGATAGAGGGGATCGCT 180
DB |||||
QY 402 TCGAAGCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGGGCGATAGAGGGGATCGCT 461
DB |||||
QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGGAGCGGTTCAGCCCATTCGCCCAA 240
DB |||||
QY 462 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGGAGCGGTTCAGCCCATTCGCCCAA 521
DB |||||
QY 241 GCTCTTCAGCAATATCACCGGTAGCCAACTATGTCTGTATAGCGGTTCGCCCAACCCA 300
DB |||||
QY 522 GCTCTTCAGCAATATCACCGGTAGCCAACTATGTCTGTATAGCGGTTCGCCCAACCCA 581
DB |||||
QY 301 GCGCGCCAGTCGATGATTCAGAAAGCGGCCATTTTCACCATGATATTCGGCAAGC 360
DB |||||
QY 582 GCGCGCCAGTCGATGATTCAGAAAGCGGCCATTTTCACCATGATATTCGGCAAGC 641
DB |||||
QY 361 AGGCATCGCATCGGTTCAGCAGAGATCTTCGCGGTTCGGGCGATCGCGCTTGGCTGG 420
DB |||||
QY 642 AGGCATCGCATCGGTTCAGCAGAGATCTTCGCGGTTCGGGCGATCGCGCTTGGCTGG 701
DB |||||
QY 421 CGAAGATTCGGGTTCGGCGGAGCCCTGATGCTTCGTCAGATCATCTCTGATCGACAA 480
DB |||||
QY 702 CGAAGATTCGGGTTCGGCGGAGCCCTGATGCTTCGTCAGATCATCTCTGATCGACAA 761
DB |||||
QY 481 GACCGGCTTCACATCCGAGTACGCTCG 508
DB |||||
QY 762 GACCGGCTTCACATCCGAGTACGCTCG 789
DB |||||

RESULT 6

AQ361914/c
LOCUS
DEFINITION mgxb0005K01f CUGI Rice Blast BAC Library Magnaporthe grisea genomic

AQ361914 789 bp DNA linear GSS 03-FEB-1999
mgxb0005K01f CUGI Rice Blast BAC Library Magnaporthe grisea genomic

```

clone mgxb0005K01f, genomic survey sequence.
ACCESSION AQ361914 GI:4211753
VERSION AQ361914.1
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
1 (bases 1 to 789)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
JOURNAL Genome
COMMENT Unpublished
CONTACT: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 41
High quality sequence stop: 392.
FEATURES
location/Qualifiers
1..789
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0005K01f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 160 a 211 c 234 g 181 t 3 others
ORIGIN
Query Match 50.6%; Score 507; DB 28; Length 789;
Best Local Similarity 99.8%; Pred. NO. 6.3e-255;
Matches 557; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 180 TCGGAATCGGAGCGCGATACCGTAAAGCAGCAGGAGCGGTGAGCCCATTCGCGGCCA 239
DB 599 TCGGAATCGGAGCGCGGTATACCGTAAAGCAGCAGGAGCGGTGAGCCCATTCGCGGCCA 540
QY 240 AGCTCTTCAGCAATATCAGGATAGCCAAACGCTATGCTCTGATAGCGGTGCGCACACC 299
DB 539 AGCTCTTCAGCAATATCAGGATAGCCAAACGCTATGCTCTGATAGCGGTGCGCACACC 480
QY 300 AGCCGGCCACACTGATGATATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAG 359
DB 479 AGCCGGCCACACTGATGATATCCAGAAAGCGGCATTTCCACCATGATATTCGGCAAG 420
QY 360 CAGGCATCCCATGGGTCAACAGAGATCTTCGCGGTGCGGCATGCGCGCTTGAGCTTG 419
DB 419 CAGGCATCCCATGGGTCAACAGAGATCTTCGCGGTGCGGCATGCGCGCTTGAGCTTG 360
QY 420 GCGAACAGTTCGCTGGCGGAGCCCTTGATGCTCTTCGTCACGATCATCTCATGACGA 479
DB 359 GCGAACAGTTCGCTGGCGGAGCCCTTGATGCTCTTCGTCACGATCATCTCATGACGA 300

```

```

QY 480 AGACCGGCTTCATCCGATACGTCGCTCGATCGATGATGTTGCTTGTGTTGCGAAT 539
DB 299 AGACCGGCTTCATCCGATACGTCGCTCGATCGATGATGTTGCTTGTGTTGCGAAT 240
QY 540 GGGCAGTAGCCGATCAAGCGTATGATGAGCGCGCATTCGATTCAGCCATGATGATAT 599
DB 239 GGGCAGTAGCCGATCAAGCGTATGATGAGCGCGCATTCGATTCAGCCATGATGATAT 180
QY 600 TTCTCGCAGAGCAGAGTGCAGATGACAGAGATCTGCGCGGCACTTCGCCCATAGC 659
DB 179 TTCTCGCAGAGCAGAGTGCAGATGACAGAGATCTGCGCGGCACTTCGCCCATAGC 120
QY 660 AGCCAGTCCCTCCGCTTCAGTGACAAAGTCGAGCAGACGTCGCGCAAGAACGCCGTC 719
DB 119 AGCCAGTCCCTCCGCTTCAGTGACAAAGTCGAGCAGACGTCGCGCAAGAACGCCGTC 60
QY 720 GTGGCCAGCCACGATAGC 737
DB 59 GTGGCCAGCCACGATAGC 42

RESULT 7
AQ398880/c
LOCUS 499 bp DNA linear GSS 06-MAR-1999
DEFINITION mgxb0006P12f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0006P12f, genomic survey sequence.
ACCESSION AQ398880
VERSION AQ398880.1 GI:4369907
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
1 (bases 1 to 499)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
JOURNAL Genome
COMMENT Unpublished
CONTACT: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence stop: 422.
FEATURES
location/Qualifiers
1..499
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0006P12f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 94 a 142 c 157 g 106 t
ORIGIN

```

```

Query Match      49.9%; Score 499; DB 28; Length 499;
Best Local Similarity 100.0%; Pred. No. 9.3e-251;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 GCCTGCGAATCGGAGCGCGATACCGTAAGACAGAGGAGCGGTACGCCCATTCGCC 235
DB 499 GCCTGCGAATCGGAGCGCGATACCGTAAGACAGAGGAGCGGTACGCCCATTCGCC 440
QY 236 GCGAAGCTTTCAGCAATATCAAGGTATGCTATGCTATGATGCGGTCCGCCAC 295
DB 439 GCGAAGCTTTCAGCAATATCAAGGTATGCTATGCTATGATGCGGTCCGCCAC 380
QY 296 ACCGAGCGCGCACATGCGATGATCCAGAAAGCGGCCATTTTCACATGATTCGG 355
DB 379 ACCGAGCGCGCACATGCGATGATCCAGAAAGCGGCCATTTTCACATGATTCGG 320
QY 356 CAAAGAGCGATCGCCATGCGGTACGACGAGATCCTCGCGTGGCGCATGCGCGCTTCGAG 415
DB 319 CAAAGAGCGATCGCCATGCGGTACGACGAGATCCTCGCGTGGCGCATGCGCGCTTCGAG 260
QY 416 CTGCGGACACATTCGCTGGCGGAGCGGCCCTGATGCTCTTCGTCAGATCATCTGATC 475
DB 259 CTGCGGACACATTCGCTGGCGGAGCGGCCCTGATGCTCTTCGTCAGATCATCTGATC 200
QY 476 GACAAGACCGGCTTCCATCCGAGTACGCTCGCTCGATGCGATGCTTTGCTTGGTGTC 535
DB 199 GACAAGACCGGCTTCCATCCGAGTACGCTCGCTCGATGCGATGCTTTGCTTGGTGTC 140
QY 536 GAATGGGAGGTAGCGCGATCAAGCGTATGAGCGCGCGCATTTGATCAGCCATGATGGA 595
DB 139 GAATGGGAGGTAGCGCGATCAAGCGTATGAGCGCGCGCATTTGATCAGCCATGATGGA 80
QY 596 TACTTCTTCGCGAGGAGCAAGGTGAGATGACAGGAGATCCTCGCGCGCATTCGCGCGCA 655
DB 79 TACTTCTTCGCGAGGAGCAAGGTGAGATGACAGGAGATCCTCGCGCGCATTCGCGCGCA 20
QY 656 TAGCGACCGATGCTTCCCTCC 674
DB 19 TAGCGACCGATGCTTCCCTCC 1

```

```

RESULT 8
LOCUS      AQ449162/c
DEFINITION mgxb00023E21f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb00023E21f, genomic survey sequence.
ACCESSION  AQ449162
VERSION     AQ449162.1 GI:4578299
KEYWORDS   GSS
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE   1 (bases 1 to 592)
AUTHORS     Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE       A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
JOURNAL     Unpublished
COMMENT     Contact: Dean RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson University, Clemson, SC 29634
            Tel: 864 656 5737
            Fax: 864 656 4293
            Email: rdean@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence stop: 395.
            Location/Qualifiers
            1..592
            /organism="Magnaporthe grisea"
            /mol_type="genomic DNA"
FEATURES
source

```

```

/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb00023E21f"
/issue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 118 a 166 c 172 g 136 t
ORIGIN

```

```

Query Match      49.9%; Score 499; DB 28; Length 592;
Best Local Similarity 100.0%; Pred. No. 9.6e-251;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CTTGCTTGGTTCGTCATTTGGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAG 161
DB 544 CTTGCTTGGTTCGTCATTTGGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAG 485
QY 162 CGATAGAAGCGGATCGGCTCGCAATCGGAGCGGGGATACCGTAAGACAGGAGGAGCGG 221
DB 484 CGATAGAAGCGGATCGGCTCGCAATCGGAGCGGGGATACCGTAAGACAGGAGGAGCGG 425
QY 222 TGAGCCCATTTGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGCTTGA 281
DB 424 TCAGCCCATTTGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGCTTGA 365
QY 282 TAGCGGTTCGCGCACACCCAGCGGCGACAGTCGATGATTCAGAAAGCGGCATTTCC 341
DB 364 TAGCGGTTCGCGCACACCCAGCGGCGACAGTCGATGATTCAGAAAGCGGCATTTCC 305
QY 342 ACCATGATATTCGGAAGCAGGATCGCCATCGGTCACGACGAGATCTTCGCGCTCGGCG 401
DB 304 ACCATGATATTCGGAAGCAGGATCGCCATCGGTCACGACGAGATCTTCGCGCTCGGCG 245
QY 402 ATGCGCGCTTCGAGCTGCGCAACAGTTGCGGTGCGGAGCCCTGATGCTCTTCGTC 461
DB 244 ATGCGCGCTTCGAGCTGCGCAACAGTTGCGGTGCGGAGCCCTGATGCTCTTCGTC 185
QY 462 AGATCATCTCTGATCGACAGACCGGCTTCCATCCGAGTACGTCGCTCGATGCGATGT 521
DB 184 AGATCATCTCTGATCGACAGACCGGCTTCCATCCGAGTACGTCGCTCGATGCGATGT 125
QY 522 TTGCGTTGGTTCGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCGA 581
DB 124 TTGCGTTGGTTCGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCGA 65
QY 582 TCAGCCATGATGATCTT 600
DB 64 TCAGCCATGATGATCTT 46

```

```

RESULT 9
LOCUS      AQ447230/c
DEFINITION mgxb00060O09f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb00060O09f, genomic survey sequence.
ACCESSION  AQ447230
VERSION     AQ447230.1 GI:4576367
KEYWORDS   GSS
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE   1 (bases 1 to 600)
AUTHORS     mgxb00060O09f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb00060O09f, genomic survey sequence.
TITLE       A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
JOURNAL     Unpublished
COMMENT     Contact: Dean RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson University, Clemson, SC 29634
            Tel: 864 656 5737
            Fax: 864 656 4293
            Email: rdean@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence stop: 395.
            Location/Qualifiers
            1..592
            /organism="Magnaporthe grisea"
            /mol_type="genomic DNA"
FEATURES
source

```

REFERENCE 1 (bases 1 to 637)
 AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinski, M., Wing, R.A. and Dean, R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
 JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 Class: BAC ends
 High quality sequence stop: 291.
 FEATURES
 Location/Qualifiers
 1..637
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0006009f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."
 127 a 170 c 187 g 152 t 1 others

BASE COUNT 127 a 170 c 187 g 152 t 1 others
 ORIGIN
 Query Match 49.3%; Score 493; DB 28; Length 637;
 Best Local Similarity 99.8%; Pred. No. 1.4e-247;
 Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 102 CGTCGCTTGTGGTCTGATTTTCGAAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGG 161
 DB 544 CGTCGCTTGTGGTCTGATTTTCGAAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGG 485
 QY 162 CGATAGAAGCGGATGCGCTCGAATCGGAGCGCGCATACCGTAAAGCACGAGGAGCGG 221
 DB 484 CGATAGAAGCGGATGCGCTCGAATCGGAGCGCGCATACCGTAAAGCACGAGGAGCGG 425
 QY 222 TCAGCCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGGTAGCCAACTATGTCTCGA 281
 DB 424 TCAGCCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGGTAGCCAACTATGTCTCGA 365
 QY 282 TAGGGTCCGCACACCCAGCGGCCACAGTCGATCAATCCAGAAAGCGGCATTTTC 341
 DB 364 TAGGGTCCGCACACCCAGCGGCCACAGTCGATCAATCCAGAAAGCGGCATTTTC 305
 QY 342 ACCATGATATTCGCGAAGCAGGCGATCGCCATGGTCCAGCAGAGATCCTCGCGCTCGGC 401
 DB 304 ACCATGATATTCGCGAAGCAGGCGATCGCCATGGTCCAGCAGAGATCCTCGCGCTCGGC 245
 QY 402 ATCGCGCCCTTAGCCTGGGCAACAGTTGGCTGGCGCGAGGCCCTGATGCTTCGTCTCC 461
 DB 244 ATCGCGCCCTTAGCCTGGGCAACAGTTGGCTGGCGCGAGGCCCTGATGCTTCGTCTCC 185
 QY 462 AGATCATCTGTATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGT 521
 DB 184 AGATCATCTGTATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGT 125
 QY 522 TTCGTTGGTGTGCAATGGGCAAGGTAGCGCGGATCAAGCGTATGCAAGCCGCCCATTTGCA 581

Db 124 TTCGTTGGTGTGCAATGGGCAAGGTAGCCGGATCAAGCGTATGAGCGCGCATTTGCA 65
 QY 582 TCAGCCCATTCGCGCGCAAGCTCTTCGCGAGGAGCAAGGTGAGATGACAGAGATCTCTGCC 641
 Db 64 TCAGCCCATTCGCGCGCAAGCTCTTCGCGAGGAGCAAGGTGAGATGACAGAGATCTCTGCC 5
 QY 642 GGCA 645
 Db 4 GGCA 1
 RESULT 10
 LOCUS AQ875040
 DEFINITION V10E4 mTn-3xHA/lacZ Insertion Library, strain Y2728 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
 ACCESSION AQ875040
 VERSION AQ875040.1 GI:6287284
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 REFERENCE 1 (bases 1 to 861)
 AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deStages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
 TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
 JOURNAL Unpublished
 COMMENT Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mTn-3xHA/lacZ insertion.
 Seq primer: GGCCTTCTTCTTTTGAAGTAC
 Class: transposon-tagged.
 FEATURES
 Location/Qualifiers
 1..861
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /strain="Y2728 - S288C background, cir(0) rho(0)"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2728"
 /note="Vector: pHSS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHSS6-Sal; Genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."
 BASE COUNT 203 a 253 c 233 g 169 t 3 others
 ORIGIN
 Query Match 49.3%; Score 493; DB 28; Length 861;
 Best Local Similarity 100.0%; Pred. No. 1.4e-247;
 Matches 493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGGAGGATCATCCAGCCGGCTCCCGAAACGATTCGGAAGCCCAACCTTTCATAG 60
 DB 215 CGCTGGAGGATCATCCAGCCGGCTCCCGAAACGATTCGGAAGCCCAACCTTTCATAG 274
 QY 61 AAGCGCGGTGGATCGAATCTCGTATGCGAGTTGGCGCTCGCTTGGTGGTCAAT 120
 DB 275 AAGCGCGGTGGATCGAATCTCGTATGCGAGTTGGCGCTCGCTTGGTGGTCAAT 334
 QY 121 TCGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAGGCGATAGAGGCGATGGCT 180


```

Db      335 TCAGACCCAGAGTCCCGCTCAGAGAACTCGTCAGAAAGCGGATAGAGGCGATGCCGT 394
QY      181 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCCGCCAA 240
Db      395 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCCGCCAA 454
QY      241 GCTCTTCAGCAATATCACGGGTAGCCAAAGCGGTATGTCCTGATAGCGGTCCGCCACACCA 300
Db      455 GCTCTTCAGCAATATCACGGGTAGCCAAAGCGGTATGTCCTGATAGCGGTCCGCCACACCA 514
QY      301 GCGGCGCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db      515 GCGGCGCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 574
QY      361 AGGCATCGGCATGGGTACAGAGAGATCCTCGCGGTCCGGGATGCGCGCTTGACCTGG 420
Db      575 AGGCATCGGCATGGGTACAGAGAGATCCTCGCGGTCCGGGATGCGCGCTTGACCTGG 634
QY      421 CGAACAGTTTCGGTGGCGGAGCCCTCGATGCTCTTCGTCAGATCATCCTGATCGACAA 480
Db      635 CGAACAGTTTCGGTGGCGGAGCCCTCGATGCTCTTCGTCAGATCATCCTGATCGACAA 694
QY      481 GACCGGCTTCAT 493
Db      695 GACCGGCTTCAT 707

```

```

RESULT 11
LOCUS   AQ876220
DEFINITION
V152G7 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
cerevisiae genomic 5', genomic survey sequence.
ACCESSION
AQ876220
VERSION
AQ876220.1 GI:6288464
KEYWORDS
GSS.
SOURCE
Saccharomyces cerevisiae (baker's yeast)
ORGANISM
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
1 (bases 1 to 799)
Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A.,
desEtages,S.A., Cheung,K.-H., Sheehan,A., Symoniatas,D., Jansen,R.,
Umansky,L., Heidtman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,
Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished
Contact: Kumar A
Location/Qualifiers
1..799
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
without 2 micron or mitochondrial DNA was prepared in
pHS6-Sal; genomic DNA was size-fractionated (DNA of
roughly 2-3 kb in length) prior to cloning. This library
was subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."

```

FEATURES
source

BASE COUNT
ORIGIN

176 a 234 c 228 g 161 t

Query Match 48.9%; Score 489; DB 28; Length 799;
Best Local Similarity 100.0%; Pred. No. 1.8e-245; Length 799;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CGTGGAGGATCATCAGCGCGCTCCCGAAACGATTCCGAAGCCCAACCTTTTCATAG 60
Db      187 CGTGGAGGATCATCAGCGCGCTCCCGAAACGATTCCGAAGCCCAACCTTTTCATAG 246
QY      61 AAGCGCGCGGTGAATCGAAATCTCGTATGCGAGGTGGCGGTTCGTTGGTTCGTCATT 120
Db      247 AAGCGCGCGGTGAATCGAAATCTCGTATGCGAGGTGGCGGTTCGTTGGTTCGTCATT 306
QY      121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAAGGCGGATCGGCT 180
Db      307 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAAGGCGGATCGGCT 366
QY      181 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCCGCCAA 240
Db      367 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCCGCCAA 426
QY      241 GCTCTTCAGCAATATCACGGGTAGCCAAAGCGGTATGTCCTGATAGCGGTCCGCCACACCA 300
Db      427 GCTCTTCAGCAATATCACGGGTAGCCAAAGCGGTATGTCCTGATAGCGGTCCGCCACACCA 486
QY      301 GCGGCGCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db      487 GCGGCGCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 546
QY      361 AGGCATCGGCATGGGTACAGAGAGATCCTCGCGGTCCGGGATGCGCGCTTGACCTGG 420
Db      547 AGGCATCGGCATGGGTACAGAGAGATCCTCGCGGTCCGGGATGCGCGCTTGACCTGG 606
QY      421 CGAACAGTTTCGGTGGCGGAGCCCTCGATGCTCTTCGTCAGATCATCCTGATCGACAA 480
Db      607 CGAACAGTTTCGGTGGCGGAGCCCTCGATGCTCTTCGTCAGATCATCCTGATCGACAA 666
QY      481 GACCGGCTTCAT 489
Db      667 GACCGGCTTCAT 675

```

RESULT 12
LOCUS AQ875907

DEFINITION
V130H3 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
cerevisiae genomic 5', genomic survey sequence.

ACCESSION
AQ875907

VERSION
AQ875907.1 GI:6288151

KEYWORDS
GSS.

SOURCE
Saccharomyces cerevisiae (baker's yeast)

ORGANISM
Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE
1 (bases 1 to 801)

AUTHORS
Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A.,

desEtages,S.A., Cheung,K.-H., Sheehan,A., Symoniatas,D., Jansen,R.,

Umansky,L., Heidtman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,

Hager,K., Miller,P., Roeder,G.S. and Snyder,M.

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and

Gene Disruption

Unpublished

Contact: Kumar A

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.

Seq primer: GGCCTTCTTTTGGGAAGTAC

FEATURES
source

Class: transposon-tagged.
Location/Qualifiers
1..801
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/laE_host="E. coli"
/clone_lib="mtn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS86-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS86-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 178 a 228 c 221 g 174 t

ORIGIN

Query Match 48.1%; Score 481; DB 28; Length 801;
Best Local Similarity 99.8%; Pred. No. 2.9e-241;
Matches 531; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGAAACGATTCGAAAGCCCAACCTTTTCATAG 60
DB 270 CGCTGGAGGATCATCCAGCGCGCTCCCGAAACGATTCGAAAGCCCAACCTTTTCATAG 329
QY 61 AAGCGCGGTGGGAATCGAATCTCGTATGCGAGGTGGCGGTGGCTTGGTCGGTCATT 120
DB 330 AAGCGCGGTGGGAATCGAATCTCGTATGCGAGGTGGCGGTGGCTTGGTCGGTCATT 389
QY 121 TCGAACCCCAAGTCCCGCTCAGAAAGATCTCGTCAAGAGCGGATAGAGGCGATCGCT 180
DB 390 TCGAACCCCAAGTCCCGCTCAGAAAGATCTCGTCAAGAGCGGATAGAGGCGATCGCT 449
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCCCAAA 240
DB 450 GCGAATCGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCCCAAA 509
QY 241 GCTCTTCAGCAATATACAGGATGAGCAACGCTATGCTGATAGCGGTCCGCCACACCA 300
DB 510 GCTCTTCAGCAATATACAGGATGAGCAACGCTATGCTGATAGCGGTCCGCCACACCA 569
QY 301 GCGGCCACAGTCAGTGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGGCAAGC 360
DB 570 GCGGCCACAGTCAGTGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGGCAAGC 629
QY 361 AGGCATCGCATGGGTACAGCAGATCTTCGCGGTGGGCATCGCGGCTTGGAGCTGG 420
DB 630 AGGCATCGCATGGGTACAGCAGATCTTCGCGGTGGGCATCGCGGCTTGGAGCTGG 689
QY 421 CGAACAGTTCGGTGGCGAGCCCTGTGATGCTCTTCTGTCAGATCATCTGATGACAA 480
DB 690 CGAACAGTTCGGTGGCGAGCCCTGTGATGCTCTTCTGTCAGATCATCTGATGACAA 749
QY 481 GACGGGTTCCATCCGAGTACGTGCTGCTGATGATGATGTTTCGTTGGTG 532
DB 750 GACGGGTTCCATCCGAGTACGTGCTGCTGATGATGATGTTTCGTTGGTG 801

RESULT 13
AQ876024
LOCUS
DEFINITION V132c9 mtn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION AQ876024
VERSION AQ876024.1
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 783)

AUTHORS

Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., desRages, S.A., Cheung, K.-H., Sheehan, A., Symoniatitis, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S., and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

TITLE

Unpublished

JOURNAL

COMMENT

Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mtn-3xHA/lacZ insertion.
Seq primer: GGCCTTCTTCTTCTTGGAAAGTAC
Class: transposon-tagged

FEATURES

source

1..783
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/laE_host="E. coli"
/clone_lib="mtn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS86-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS86-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 181 a 224 c 223 g 154 t 1 others

ORIGIN

Query Match 47.9%; Score 479; DB 28; Length 783;
Best Local Similarity 100.0%; Pred. No. 3.2e-240;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGAAACGATTCGAAAGCCCAACCTTTTCATAG 60
DB 296 CGCTGGAGGATCATCCAGCGCGCTCCCGAAACGATTCGAAAGCCCAACCTTTTCATAG 355
QY 61 AAGCGCGGTGGGAATCGAATCTCGTATGCGAGGTGGCGGTGGCTTGGTCGGTCATT 120
DB 356 AAGCGCGGTGGGAATCGAATCTCGTATGCGAGGTGGCGGTGGCTTGGTCGGTCATT 415
QY 121 TCGAACCCCAAGTCCCGCTCAGAAAGATCTCGTCAAGAGCGGATAGAGGCGATCGCT 180
DB 416 TCGAACCCCAAGTCCCGCTCAGAAAGATCTCGTCAAGAGCGGATAGAGGCGATCGCT 475
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCCCAAA 240
DB 476 GCGAATCGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCCCAAA 535
QY 241 GCTCTTCAGCAATATACAGGATGAGCAACGCTATGCTGATAGCGGTCCGCCACACCA 300
DB 536 GCTCTTCAGCAATATACAGGATGAGCAACGCTATGCTGATAGCGGTCCGCCACACCA 595
QY 301 GCGGCCACAGTCAGTGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGGCAAGC 360
DB 596 GCGGCCACAGTCAGTGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGGCAAGC 655
QY 361 AGGCATCGCATGGGTACAGCAGATCTTCGCGGTGGGCATCGCGGCTTGGAGCTGG 420
DB 656 AGGCATCGCATGGGTACAGCAGATCTTCGCGGTGGGCATCGCGGCTTGGAGCTGG 715
QY 421 CGAACAGTTCGGTGGCGAGCCCTGTGATGCTCTTCTGTCAGATCATCTGATGACAA 479
DB 716 CGAACAGTTCGGTGGCGAGCCCTGTGATGCTCTTCTGTCAGATCATCTGATGACAA 774

RESULT 14

```

A0447874/c
LOCUS          A0447874          561 bp      DNA          linear      GSS 08-APR-1999
DEFINITION     mgxb0012101f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
                clone mgxb0012101f, genomic survey sequence.
ACCESSION      A0447874
VERSION        A0447874.1  GI:4577011
KEYWORDS       GSS.
SOURCE         Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM       Magnaporthe grisea
REFERENCE      1 (bases 1 to 561)
AUTHORS        Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
                Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE          A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
                Genome
JOURNAL        Unpublished
COMMENT        Contact: Dean RA
                Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson University, Clemson, SC 29634
                Tel: 864 656 5737
                Fax: 864 656 4293
                Email: rdean@clemson.edu
                Seq primer: TAATACGACTCACTATAGGG
                Class: BAC ends
                High quality sequence start: 42
                High quality sequence stop: 326.
                Location/Qualifiers
FEATURES       1..561
                /organism="Magnaporthe grisea"
                /mol_type="genomic DNA"
                /strain="70-15"
                /db_xref="taxon:148305"
                /clone="mgxb0012101f"
                /tissue_type="Protoplasts"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice Blast BAC Library"
                /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
                Rice blast is one of the most devastating fungal diseases
                of rice world wide. It is a filamentous ascomycete with
                a haploid genome (n=7) of approximately 40 Mbp. Rice
                blast is an important model fungal pathogen for studying
                numerous aspects of the fungal-host interaction. In
                order to facilitate genome wide analysis, a BAC library
                containing 9216 clones with an average insert size of 130
                kbp was constructed. This library represents greater
                than 25X genome coverage. High density colony filters
                are available upon request."
BASE COUNT     104 a 163 c 174 g 118 t
ORIGIN
Query Match    46.8%; Score 468; DB 28; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.8e-234;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 294 ACACCCAGCGGCGCACAGTCGATCGATTCAGAAAGCGGCATTTCCACCATGATTC 353
Db 561 ACACCCAGCGGCGCACAGTCGATCGATTCAGAAAGCGGCATTTCCACCATGATTC 502
Qy 354 GGCAAGCAGGCATCGCATGGTTCACGACGAGATCTTCGCGCGGCATGCGCGCTTG 413
Db 501 GGCAGCAGGCATCGCATGGTTCACGACGAGATCTTCGCGCGGCATGCGCGCTTG 442
Qy 414 AGCCTGCGCAACAGTTCGGCTGGCGGAGCCCTTGATGCTTCTTCGTCAGATCATCTCGA 473
Db 441 AGCCTGCGCAACAGTTCGGCTGGCGGAGCCCTTGATGCTTCTTCGTCAGATCATCTCGA 382
Qy 474 TCGACAGACGGCTTCATCCGAGTAGTGTCTCGCTCGATCGATGTTTCGTTGGTG 533
Db 381 TCGACAGACGGCTTCATCCGAGTAGTGTCTCGCTCGATCGATGTTTCGTTGGTG 322
Qy 534 TCGAATGGCAGGTAGCGCGATCAAGCGTATGACGCGCGCGCATTCGATCAGCCATGATG 593

```

```

321 TCGAATGGCAGGTAGCGGATCAAGCGTATGTCAGCCGCCCGCATTCATCATGCCATGATG 262
Qy 594 GATCTTTCTCGGAGGAGCAAGGTGAGTACACAGGAGATCCTGCCGCGGCACTTCGCCC 653
Db 261 GATCTTTCTCGGAGGAGCAAGGTGAGTACACAGGAGATCCTGCCGCGGCACTTCGCCC 202
Qy 654 AATAGCAGCCAGTCCCTTCCCGCTTCAGTTCAGTACAACTCGAGCAGCTGGCCAGGAACG 713
Db 201 AATAGCAGCCAGTCCCTTCCCGCTTCAGTTCAGTACAACTCGAGCAGCTGGCCAGGAACG 142
Qy 714 CCCGTCGTCGCGCAGCAGCAGTAGCCGCGCTGCCTTCCTTCGTCAGTTCA 761
Db 141 CCCGTCGTCGCGCAGCAGCAGTAGCCGCGCTGCCTTCCTTCGTCAGTTCA 94

RESULT 15
A0398387/c
LOCUS          A0398387          620 bp      DNA          linear      GSS 06-MAR-1999
DEFINITION     mgxb0009005f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
                clone mgxb0009005f, genomic survey sequence.
ACCESSION      A0398387
VERSION        A0398387.1  GI:4369414
KEYWORDS       GSS.
SOURCE         Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM       Magnaporthe grisea
REFERENCE      1 (bases 1 to 620)
AUTHORS        Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
                Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE          A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
                Genome
JOURNAL        Unpublished
COMMENT        Contact: Dean RA
                Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson University, Clemson, SC 29634
                Tel: 864 656 5737
                Fax: 864 656 4293
                Email: rdean@clemson.edu
                Seq primer: TAATACGACTCACTATAGGG
                Class: BAC ends
                High quality sequence stop: 187.
                Location/Qualifiers
FEATURES       1..620
                /organism="Magnaporthe grisea"
                /mol_type="genomic DNA"
                /strain="70-15"
                /db_xref="taxon:148305"
                /clone="mgxb0009005f"
                /tissue_type="Protoplasts"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice Blast BAC Library"
                /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
                Rice blast is one of the most devastating fungal diseases
                of rice world wide. It is a filamentous ascomycete with
                a haploid genome (n=7) of approximately 40 Mbp. Rice
                blast is an important model fungal pathogen for studying
                numerous aspects of the fungal-host interaction. In
                order to facilitate genome wide analysis, a BAC library
                containing 9216 clones with an average insert size of 130
                kbp was constructed. This library represents greater
                than 25X genome coverage. High density colony filters
                are available upon request."
BASE COUNT     121 a 172 c 188 g 139 t
ORIGIN
Query Match    46.8%; Score 468; DB 28; Length 620;
Best Local Similarity 100.0%; Pred. No. 1.9e-234;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 245 TTCAGCAATATCAGCGGTAGCCCAACGCTATGTCCTGTATAGCGGTCCGCCACACCCAGCG 304

```

```
Db 468 TTCCAGCAATATCACGGTAGCCAAAGCTATATGCTCTGATAGCGGTCCGCCACACCCAGCGG 409
Qy 305 GCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGC 364
Db 408 GCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGC 349
Qy 365 ATCGCCATGGGTACAGACGAGATCCTCGCCGTCCGGGCATGCGCGCTTGAGCCTGGCGAA 424
Db 348 ATCGCCATGGGTACAGACGAGATCCTCGCCGTCCGGGCATGCGCGCTTGAGCCTGGCGAA 289
Qy 425 CAGTTCCGCTGGCGAGAGCCCTGATGCTTTCGTCAGATCATCCTGATCGACAAGACC 484
Db 288 CAGTTCCGCTGGCGAGAGCCCTGATGCTTTCGTCAGATCATCCTGATCGACAAGACC 229
Qy 485 GGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTTCGTTGGTGGTCCGATGGCA 544
Db 228 GGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTTCGTTGGTGGTCCGATGGCA 169
Qy 545 GTAGCCGGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGGATATTTTCTC 604
Db 168 GTAGCCGGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGGATATTTTCTC 109
Qy 605 GGCAGGAGCAAGGTGAGATGACAGAGATCCTGCGCCCGGCATTCGCCCAATAGCAGCCA 664
Db 108 GGCAGGAGCAAGGTGAGATGACAGAGATCCTGCGCCCGGCATTCGCCCAATAGCAGCCA 49
Qy 665 GTCCTTCCCGCTTCAGTGACCAACGTCCGAGCACAGCTCGCAAGAAC 712
Db 48 GTCCTTCCCGCTTCAGTGACCAACGTCCGAGCACAGCTCGCAAGAAC 1
```

Search completed: February 16, 2004, 03:15:50
Job time : 2641.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 25, 2004, 16:56:30 ; Search time 6417 Seconds
(without alignments)
11306.878 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gtccttcaccatcactcg.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.pl.*
- 8: gb.pr.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1674	100.0	1674	6	BD270607	BD270607 Vascular
2	1674	100.0	1674	6	AR381510	AR381510 Sequence
3	1674	100.0	1674	6	BD001745	BD001745 Human vas
4	1674	100.0	1674	6	BD134149	BD134149 Vascular
5	1666	99.5	1948	9	BC035212	BC035212 Homo sapi
6	1664.6	99.4	2031	6	AR231261	AR231261 Sequence
7	1664.6	99.4	2031	6	AR342826	AR342826 Sequence
8	1664.6	99.4	2031	6	AR342826	AR342826 Sequence
9	1663	99.3	2031	6	AR231262	AR231262 Homo sapi
10	1663	99.3	2031	6	AR342827	AR342827 Sequence
11	1651.8	98.7	1997	6	AR112510	AR112510 Sequence
12	1651.8	98.7	1997	6	AR157589	AR157589 Sequence
13	1651.8	98.7	1997	6	BD247178	BD247178 Use of VE
14	1651.8	98.7	1997	6	AR201986	AR201986 Sequence
15	1651.8	98.7	1997	6	AR212635	AR212635 Sequence
16	1651.8	98.7	1997	6	AR430033	AR430033 Sequence
17	1651.8	98.7	1997	6	AX234371	AX234371 Sequence
18	1651.8	98.7	1997	6	AX743106	AX743106 Sequence
19	1651.8	98.7	1997	6	BD082159	BD082159 Vascular
20	1651.8	98.7	1997	6	HSVEGFC	X94216 H.sapiens m
21	1651.8	98.7	2015	9	HSU43142	U43142 Human vascu
22	1651.8	98.7	2679	6	BD247192	BD247192 Use of VE
23	1620.6	96.8	1939	9	HSU58111	U58111 Human FLT4
24	1526	91.2	1526	6	BD270608	BD270608 Vascular
25	1526	91.2	1526	6	BD134150	BD134150 Vascular
26	1292.6	77.2	1777	4	AB004275	AB004275 Bos tauru
27	1256.8	75.1	1260	6	AX481509	AX481509 Sequence
28	1212.4	72.4	1804	10	MMU58112	U58112 Mus musculu
29	1182.2	70.6	1818	10	MMU73620	U73620 Mus musculu
30	1182.2	70.6	1836	6	AR112511	AR112511 Sequence
31	1182.2	70.6	1836	6	AR201987	AR201987 Sequence
32	1182.2	70.6	1836	6	AR212641	AR212641 Sequence
33	1182.2	70.6	1836	6	AR430039	AR430039 Sequence
34	1182.2	70.6	1836	6	BD082160	BD082160 Vascular
35	1138.4	68.0	1140	6	AR157586	AR157586 Sequence
36	1076.4	64.3	1596	10	AY032729	AY032729 Rattus no
37	860.8	51.4	1218	10	AF432867	AF432867 Meriones
38	805.6	48.1	1257	5	CCV15937	Y15937 Coturnix co
39	804.2	48.0	1741	6	AR112512	AR112512 Sequence
40	804.2	48.0	1741	6	AR201988	AR201988 Sequence
41	804.2	48.0	1741	6	BD082161	BD082161 Vascular
42	486.4	29.1	126929	9	AC093801	AC093801 Homo sapi
43	381	22.8	1936	5	AF466147	AF466147 Danilo rer
44	335.6	20.0	177771	2	AC140905	AC140905 Homo sapi
45	321	19.2	378	4	AF009178	AF009178 Bos tauru

ALIGNMENTS

RESULT 1	BD270607	1674 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD270607	Vascular endothelial growth factor 2.			
DEFINITION	BD270607				
ACCESSION	BD270607.1	GI:33080375			
VERSION	JP 2002539082-A/1.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1674)				
AUTHORS	Rosen, C.A., Alderson, R., Melder, R., Roschke, V. and Ruben, S.M.				
TITLE	Vascular endothelial growth factor 2				
JOURNAL	Patent: JP 2002539082-A 1 19-NOV-2002;				

HUMAN GENOME SCIENCES INC		
OS	Homo sapiens (human)	
PN	JP 2002539082-A/1	
PD	19-NOV-2002	
PF	07-FEB-2000 JP 2000596954	
PR	08-FEB-1999 US 60/119179, 12-FEB-1999 US 60/119926 PR	
03-JUN-1999 US 60/137796, 22-DEC-1999 US 60/171505 FI	CRAIG	
A. ROSEN, RALPH ALDERSON, ROBERT WELDER, VIKTOR ROSCHKE, PI	STEVEN M	
RUBEN		
PC	A61K38/22, A61K38/00, A61K38/21, A61K38/27, A61K39/395, A61K39/395,	
PC	A61K45/00,	
PC	A61K47/48, A61K48/00, A61P3/10, A61P9/10, A61P25/02, A61P27/02, PC	
PC	A61P29/00,	
PC	A61P31/12, A61P37/00, C12N5/10//A61K35/12, C07K16/18, C12N15/09,	
PC	A61K37/24,	
PC	C12N5/00, A61K37/02, A61K37/66, A61K37/36, C12N15/00 CC	Vascular
endothelial growth factor 2		
EH Key	Location/Qualifiers	
FT CDS	(12)..(1268)	
FT sig peptide	(12)..(80)	
FT mat peptide	(81)..(1268)	
FEATURES	Location/Qualifiers	
source	1..1674	
ORIGIN	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
Query March	100.0%; Score 1674; DB 6; Length 1674;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1674; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GTCTTCCACCATGCACTGCTGGGCTTCTTCTGTCGGCTGTTCTCTGCTGCGCGTG 60
Db	1	GTCTTCCACCATGCACTGCTGGGCTTCTTCTGTCGGCTGTTCTCTGCTGCGCGTG 60
QY	61	CGCTGCTCCGGGTCTCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
Db	61	CGCTGCTCCGGGTCTCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
QY	121	ACCTCTCGGACGGGAGCCGCGCGGGGAGCGCCACGSCCTATGCGAGCAAGACATCTGG 180
Db	121	ACCTCTCGGACGGGAGCCGCGCGGGGAGCGCCACGSCCTATGCGAGCAAGACATCTGG 180
QY	181	AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTCTACCCAGAAT 240
Db	181	AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTCTACCCAGAAT 240
QY	241	ATTGGAATGTACAAAGTGTACAGTGTAAAGGAAGGAGGCTGCGCAATACAGAGAAACAG 300
Db	241	ATTGGAATGTACAAAGTGTACAGTGTAAAGGAAGGAGGCTGCGCAATACAGAGAAACAG 300
QY	301	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAGCATTTATAATACAG 360
Db	301	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAGCATTTATAATACAG 360
QY	361	AGATCTTGAAGTATTCGATATGATGTCGGAAGAGACTCAATGATGCGAGGAGGTGT 420
Db	361	AGATCTTGAAGTATTCGATATGATGTCGGAAGAGACTCAATGATGCGAGGAGGTGT 420
QY	421	GTATAGATGTGGGAGGAGTGTGGAGTCGGCAAAACACCTTCTTTAAACCTCCATGTG 480
Db	421	GTATAGATGTGGGAGGAGTGTGGAGTCGGCAAAACACCTTCTTTAAACCTCCATGTG 480
QY	481	TGTCGCTCTACAGATGTGGGGTGTCTGCAATAGTAGGAGGCTGCATGCGAGGAGGTGT 540
Db	481	TGTCGCTCTACAGATGTGGGGTGTCTGCAATAGTAGGAGGCTGCATGCGAGGAGGTGT 540
QY	541	GCACGAGTCTACGAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Db	541	GCACGAGTCTACGAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
QY	601	ACCAGTAAATCATGTTTGGCAATCATCACTTCTCTCGGATGCAATGCTTAAACTGGATG 660

601	AA	CCAGTAA	CAAT	CAGT	TTT	GC	CAAT	CA	CAC	TT	CT	GC	CG	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	AT	CG	
-----	----	---------	------	------	-----	----	------	----	-----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	--

RESULT 2

AR381510
LOCUS AR381510 1674 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6608182.
ACCESSION AR381510
VERSION AR381510.1 GI:40089598
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1674)
AUTHORS Rosen, C., Hu, J.-S. and Cao, L.
TITLE Human vascular endothelial growth factor 2
JOURNAL Patent: US 6608182-A 1 19-AUG-2003;
FEATURES
Location/Qualifiers
1..1674
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1674; DB 6; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCCTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGGGTGTCTCTCTCGCTCGCGCTG 60
Db 1 GTCCTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGGGTGTCTCTCTCGCTCGCGCTG 60
Qy 61 CGTGTCTCCGGGTCTCTCGAGGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Db 61 CGTGTCTCCGGGTCTCTCGAGGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Qy 121 ACCTCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGAGATCTGG 180
Db 121 ACCTCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGAGATCTGG 180
Qy 181 AGGAGCAGTTACGGTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 241 ATTGGAAATGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 ATTGGAAATGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Qy 301 CCAACCTCAACTCAAGGACAGAGACTATAAATTTGCTGACGACATTTATAATACAG 360
Db 301 CCAACCTCAACTCAAGGACAGAGACTATAAATTTGCTGACGACATTTATAATACAG 360
Qy 361 AGATCTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGATGCTCCAGGAGGTG 420
Db 361 AGATCTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGATGCTCCAGGAGGTG 420
Qy 421 GTATAGATGTGGGAGAGGTTTGGAGTGGCGACAAACCTCTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGAGAGGTTTGGAGTGGCGACAAACCTCTTTAAACCTCCATGTG 480
Qy 481 TGTCCGCTACAGATGTGGGGTGTCTGCAATAGTGGAGGCTCGAGTGAATGAACCA 540
Db 481 TGTCCGCTACAGATGTGGGGTGTCTGCAATAGTGGAGGCTCGAGTGAATGAACCA 540
Qy 541 GCAGAGCTACTCAGCAGAGAGTTTGAATTAAGTACGTCTCTCTCAAGGCCCA 600
Db 541 GCAGAGCTACTCAGCAGAGAGTTTGAATTAAGTACGTCTCTCTCAAGGCCCA 600
Qy 601 AACCAAGTAACTCAAGTTTGGCAATCACACTTCTGCGGATGATGCTAAACTGGATG 660
Db 601 AACCAAGTAACTCAAGTTTGGCAATCACACTTCTGCGGATGATGCTAAACTGGATG 660
Qy 661 TTATCAGACAGTTCAATTCATTTAGACGTTTCTGCGGAGCAACACTACCAAGTGT 720
Db 661 TTATCAGACAGTTCAATTCATTTAGACGTTTCTGCGGAGCAACACTACCAAGTGT 720
Qy 721 AGGCAGCGAAGACAGCTGCCCAATTTACATGTGGAAATACATCTGCAAGATGCC 780
Db 721 AGGCAGCGAAGACAGCTGCCCAATTTACATGTGGAAATACATCTGCAAGATGCC 780

Db 721 AGGCAGCGAAGACAGCTGCCCAATTTACATGTGGAAATACATCTGCAAGATGCC 780
Qy 781 TGGCTCAGAGAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGATTC 840
Db 781 TGGCTCAGAGAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGATTC 840
Qy 841 ATGACATCTGTGGACCAACAGAGGCTGGATGAGAGACCTGTCAGTGTCTGTCAGAG 900
Db 841 ATGACATCTGTGGACCAACAGAGGCTGGATGAGAGACCTGTCAGTGTCTGTCAGAG 900
Qy 901 CGGGCTTTCGGCTCGCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 901 CGGGCTTTCGGCTCGCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Qy 961 GTGTCTGTAATAAAACAACTCTTCCCGAGCCTTCCCGAGCCTTCCCGAGCCTTCCCGAGC 1020
Db 961 GTGTCTGTAATAAAACAACTCTTCCCGAGCCTTCCCGAGCCTTCCCGAGCCTTCCCGAGC 1020
Qy 1021 ACACATGCCAGTGTGTATGTAAGAAAGCTGCCCGCAGAAATCAACCCCTAAATCTCGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAGAAAGCTGCCCGCAGAAATCAACCCCTAAATCTCGAA 1080
Qy 1081 AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTC 1140
Db 1081 AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTC 1140
Qy 1141 ACCACCAACATGCACTGTGTACAGCGCCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAACATGCACTGTGTACAGCGCCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Qy 1261 AATGAGCTAAGATGTACTGT 1320
Db 1261 AATGAGCTAAGATGTACTGT 1320
Qy 1321 TGCCACAGTAGAATCTGTGTGAAACAGAGAGACCTTGTGGTCCATGCTAAACAAGACA 1380
Db 1321 TGCCACAGTAGAATCTGTGTGAAACAGAGAGACCTTGTGGTCCATGCTAAACAAGACA 1380
Qy 1381 AAAGTCTGTCTTCTGTAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 1381 AAAGTCTGTCTTCTGTAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Qy 1441 CAAAAGGCTCTGTGAAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 1441 CAAAAGGCTCTGTGAAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Qy 1501 TTGTGATTTCTTTAAAGAACTGATATATATATATATATATATATATATATATATATATAT 1560
Db 1501 TTGTGATTTCTTTAAAGAACTGATATATATATATATATATATATATATATATATATATAT 1560
Qy 1561 ATTCAATTTATAGCAACCAATTTGGTAAACTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 1561 ATTCAATTTATAGCAACCAATTTGGTAAACTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Qy 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTATATAAATAAATAAATAAATAAATA 1674
Db 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTATATAAATAAATAAATAAATAAATA 1674

RESULT 3
BD001745 1674 bp DNA linear PAT 31-JAN-2002
LOCUS Human vascular endothelial growth factor 2.
DEFINITION Human vascular endothelial growth factor 2.
ACCESSION BD001745
VERSION BD001745.1 GI:18626304
KEYWORDS JP 2000069982-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

LOCUS BD134149 1674 bp DNA linear PAT 18-SEP-2002
 DEFINITION Vascular endothelial growth factor 2.
 ACCESSION BD134149
 VERSION BD134149.1 GI:23229094
 KEYWORDS JP 2002505873-A/1.
 SOURCE unidentified
 ORGANISM unclassified.
 1 (bases 1 to 1674)
 REFERENCE Rosen,C.A., Kao,L. and Fu,J.S.
 AUTHORS Vascular endothelial growth factor 2
 TITLE Patent: JP 2002505873-A 1 26-FEB-2002;
 JOURNAL HUMAN GENOME SCIENCES INC
 COMMENT OS Unidentified
 PN JP 2002505873-A/1
 PD 26-FEB-2002
 PF 10-MAR-1999 JP 2000535731
 PR 13-MAR-1998 US 09/042105, 30-JUN-1998 US 09/107997 PI
 CRAIG A. ROSEN, LIANG KAO, JIN SHAN FU
 PC C12N1/21, A61K38/22, A61K48/00, A61P9/10, A61P17/06, PC
 A61P19/02,
 PC C07K14/50, C07K16/22, C12N5/10, C12N15/09, C12Q1/68, C12Q1/70, PC
 A61K37/24,
 PC C12N5/00, C12N15/00
 CC Strandedness: Double;
 CC Topology: Linear;
 CC Vascular endothelial growth factor 2
 FH Key Location/Qualifiers
 FT sig_peptide 12..80
 FT mat_peptide 81..1268
 FT CDS 12..1268.
 FEATURES
 source
 1..1674
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 1674; DB 6; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTGTCTGCTCGCGGTG 60
 DB 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTGTCTGCTCGCGGTG 60
 QY 61 CGTGCTCCGGGTCTCTCGAGGGGCCCGCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120
 DB 61 CGTGCTCCGGGTCTCTCGAGGGGCCCGCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120
 QY 121 ACCTCTCGGACCGGAGCCGCGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 121 ACCTCTCGGACCGGAGCCGCGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 181 AGGACAGTTACGGTCTGTGTCCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAA 240
 DB 181 AGGACAGTTACGGTCTGTGTCCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAA 240
 QY 241 ATTGGAAATGTACAAGTGTAGCTAAGGAAAGAGGCTGGCAATACAGAGAACAGG 300
 DB 241 ATTGGAAATGTACAAGTGTAGCTAAGGAAAGAGGCTGGCAATACAGAGAACAGG 300
 QY 301 CCAACTCAACTCAAGGACAGAGAGACTATATAAATTTGCTGCAGCACATTATATACAG 360
 DB 301 CCAACTCAACTCAAGGACAGAGAGACTATATAAATTTGCTGCAGCACATTATATACAG 360
 QY 361 AGATCTTGAAGATGTATATAGTGGAGAAAGACTCAATGATGCCCGGAGGTGT 420
 DB 361 AGATCTTGAAGATGTATATAGTGGAGAAAGACTCAATGATGCCCGGAGGTGT 420
 QY 421 GTATAGATGTGGGAGGAGTTTGAGTCCGCAACACACCTCTTTAAACCTCATGTG 480
 DB 421 GTATAGATGTGGGAGGAGTTTGAGTCCGCAACACACCTCTTTAAACCTCATGTG 480

QY 481 TGTCCGTCTACAGATGTGGGGTTCCTGCAATAGTAGAGGGGCTGCAGTGCATGAACACCA 540
 DB 481 TGTCCGTCTACAGATGTGGGGTTCCTGCAATAGTAGAGGGGCTGCAGTGCATGAACACCA 540
 QY 541 GCACGAGCTACCTCAGCAAGACGTTATTGTAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
 DB 541 GCACGAGCTACCTCAGCAAGACGTTATTGTAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
 QY 601 AACAGTAACAAATCAGTTTTTGGCCAAATACACTTCTCGCGATGATGTCTAACTGGATG 660
 DB 601 AACAGTAACAAATCAGTTTTTGGCCAAATACACTTCTCGCGATGATGTCTAACTGGATG 660
 QY 661 TTACAGACCAAGTTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACACAGTGTGC 720
 DB 661 TTACAGACCAAGTTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACACAGTGTGC 720
 QY 721 AGGACGCAACAAAGCTGCCCCCAATTAATCATGTGGAATTAATCATCTGSCAGATGCC 780
 DB 721 AGGACGCAACAAAGCTGCCCCCAATTAATCATGTGGAATTAATCATCTGSCAGATGCC 780
 QY 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATGCC 840
 DB 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATGCC 840
 QY 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTGCTGTGTCTGTGAGAG 900
 DB 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTGCTGTGTGTGCTGAGAG 900
 QY 901 CGGGCTTCGGCCTGCGAGCTGTGGACCCCAAAAGAACTAGACAGAAACTCATGCCAGT 960
 DB 901 CGGGCTTCGGCCTGCGAGCTGTGGACCCCAAAAGAACTAGACAGAAACTCATGCCAGT 960
 QY 961 GTGTCTGTAAACCAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
 DB 961 GTGTCTGTAAACCAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
 QY 1021 ACATGCCAGTGTGTATGTAAAGAACTCTGCCCGAGAAATCAACCCCTAAATCTCTGAAA 1080
 DB 1021 ACATGCCAGTGTGTATGTAAAGAACTCTGCCCGAGAAATCAACCCCTAAATCTCTGAAA 1080
 QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTTC 1140
 DB 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTTC 1140
 QY 1141 ACCACCAAACTGAGCTGTACAGCGGCATGTACAAACCGCCAGAGGCTTTGTGAGC 1200
 DB 1141 ACCACCAAACTGAGCTGTACAGCGGCATGTACAAACCGCCAGAGGCTTTGTGAGC 1200
 QY 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGCTGCTTCCCTTCATATGTCAGAAAGACCA 1260
 DB 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGCTGCTTCCCTTCATATGTCAGAAAGACCA 1260
 QY 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTCTTATTATGGAACCTGTGT 1320
 DB 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTCTTATTATGGAACCTGTGT 1320
 QY 1321 TGGCCACAGTAGAAGTGTGTGTGAACAGAGAGACCTTTGTGGGTCCATGCTTAAACAAAGACA 1380
 DB 1321 TGGCCACAGTAGAAGTGTGTGTGAACAGAGAGACCTTTGTGGGTCCATGCTTAAACAAAGACA 1380
 QY 1381 AAGAGTCTGCTTTCCTGAAACCATGTGGATAACTTTTACAGAAATGGAGCTGAGCTCATCTG 1440
 DB 1381 AAGAGTCTGCTTTCCTGAAACCATGTGGATAACTTTTACAGAAATGGAGCTGAGCTCATCTG 1440
 QY 1441 CAAAGGCTCTTTGTAAGAGCTGTTTCTGCGCAATGACCAACACGACCAAGATTTTCTCTC 1500
 DB 1441 CAAAGGCTCTTTGTAAGAGCTGTTTCTGCGCAATGACCAACACGACCAAGATTTTCTCTC 1500
 QY 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGTC 1560
 DB 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGTC 1560

```
QY 1561 ATTCAATTTTATAGCACACCAATTCGTAAACCTCCTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCACACCAATTCGTAAACCTCCTGATCAATATTTTATATCAT 1620

QY 1621 GCAAAATATGTTTAAATATAAATGAAATTTGTTATTTATATAAAAAAAAAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATATAAATGAAATTTGTTATTTATATAAAAAAAAAAAAAA 1674

RESULT 5
BC035212 1948 bp mRNA linear PRI 07-OCT-2003
LOCUS Homo sapiens vascular endothelial growth factor C, mRNA (cdna clone
DEFINITION MGC:24887 IMAGE:4753457), complete cds.
ACCESSION BC035212
VERSION BC035212.1 GI:23273809
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1948)
AUTHORS Srauszberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,
Altschul,S.F., Zebberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Fakher,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.C., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
MEDLINE 12477932
PUBMED
REFERENCE 2 (bases 1 to 1948)
AUTHORS Srauszberg,R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: ang@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 29 Row: K Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19924300.
Location/Qualifiers
```

```
source 1. 1948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:24887 IMAGE:4753457"
/tissue_type="Skin, squamous cell carcinoma"
/clone_id="NCI_CGAP_Skn4"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6.1"
1. 1948
/gene="VEGFC"
/notes="synonyms: VRP, Flt4-L"
/db_xref="LocusID:7424"
/db_xref="MIM:601528"
266..1525
/codon_start=1
/product="vascular endothelial growth factor C
preproprotein"
/protein_id="AAH35212.1"
/db_xref="GI:23273810"
/db_xref="LocusID:7424"
/translation="MHLGGFVSACLLAALLPGPREAPAAAAAPEGDLSDARPDA
AGATAVAKDLSEQLRSVSSVDLMTLVPEYKMYKCOLRGGWHQHNREQLNSR
TEETIKFAAAHYNTIELKSIDNEWRTQCMPEVCIQVKGKFGVATNFFKPCVSVY
RCGCCNSEGLQWNTSTYLSKNTFTVPLSQGPKPTISFVANTSCRMGKLDVY
RQVHSIRRLSLPATLPOCAANTCTPTVMNNHICRLAIEDFMSSDAGDSDTDFG
HDLCGPNKIDESTCCQVCRAGLRPASGPKELDRNSOCVCKNKLFPSCCCANREP
DENTCCQVCKETCPNQPLNPKACACECTESPKCLLKGGKFFHQTCVYRRPCTNRQ
KAEPPGSYSSEVRCVPVSWKRPQMS"
650..904
/misc_feature /note="PDGF; Region: Platelet-derived and vascular
endothelial growth factors (PDGF, VEGF) family"
/db_xref="CDD:smart00141"

ORIGIN
Query Match 99.5%; Score 1666; DB 9; Length 1948;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1669; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCTTCACCATGCACTCGCTGGGCTTCTCTGTGGCGGTGTCTCTGTCGCCCGCTG 60
Db 255 GTCTTTCCATGCACTTCTGCTGGGCTTCTCTGTGGCGGTGTCTCTGTCGCCCGCTG 314

QY 61 CGTGTCTCCGGGTCTCTGCGAGCGCGCCGCGCGCGCGCTTCAGTTCGGACTCG 120
Db 315 CGTGTCTCCGGGTCTCTGCGAGCGCGCGCGCGCGCGCTTCAGTTCGGACTCG 374

QY 121 ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCTTCAGTTCGGACTCG 180
Db 375 ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCTTCAGTTCGGACTCG 434

QY 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTGTACTTACCCGAAT 240
Db 435 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTGTACTTACCCGAAT 494

QY 241 ATTGGAAAATGTACAGTGTACGTAGGAAAGAGCGCTGGCAACATACAGAGACAG 300
Db 495 ATTGGAAAATGTACAGTGTACGTAGGAAAGAGCGCTGGCAACATACAGAGACAG 554

QY 301 CCAACCTCACTCAAGGACAGAGACTATAAAATTTGCTGAGCAGACATTATAATACAG 360
Db 555 CCAACCTCACTCAAGGACAGAGACTATAAAATTTGCTGAGCAGACATTATAATACAG 614

QY 361 AGATTTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGCATGCCAGGAGGTGT 420
Db 615 AGATTTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGCATGCCAGGAGGTGT 674

QY 421 GTATAGATGTGGGAGAGGAGTTGGAGTCGCGCAACACACCTTCTTTAAACCTCCATGTG 480
Db 675 GTATAGATGTGGGAGAGGAGTTGGAGTCGCGCAACACACCTTCTTTAAACCTCCATGTG 734

QY 481 TGTCCCGTACAGATGTGGGGGTGTGTCGAATAGTGAGGGGTGTCAGTGCATGAACACCA 540
```

Db 735 TGTCGGTCTACAGATGTGGGGTCTGCAATAGTAGGGGCTGACGTGCAATGACACCA 794
Qy 541 GCACAGCTACTCTACAGAGAGCTTATTTGAAATACAGTCCCTCTCTCAAGGCCCA 600
Db 795 GCACAGCTACTCTACAGAGAGCTTATTTGAAATACAGTCCCTCTCTCAAGGCCCA 854
Qy 601 AACAGTAAATCAGTTTTCCTCAATACACTCTCTGCGATGATGCTAACTGGATG 660
Db 855 AACAGTAAATCAGTTTTCCTCAATACACTCTCTGCGATGATGCTAACTGGATG 914
Qy 661 TTACAGACAGTTCATTCCTCAATATAGACCTTCCCTGCCAGCAACACTACCAAGTGC 720
Db 915 TTACAGACAGTTCATTCCTCAATATAGACCTTCCCTGCCAGCAACACTACCAAGTGC 974
Qy 721 AGGAGCGAACAGACCTGCCCAATACATGCTGAATAATCACTCTGAGATGCC 780
Db 975 AGGAGCGAACAGACCTGCCCAATACATGCTGAATAATCACTCTGAGATGCC 1034
Qy 781 TGGCTCAGAGATTTATTTATTTTCTCGGATGCTGGAGATGACTCAACAGATGCC 840
Db 1035 TGGCTCAGAGATTTATTTATTTTCTCGGATGCTGGAGATGACTCAACAGATGCC 1094
Qy 841 ATGACATCTGGACCAACAGAGCTGGATGAGAGACTGTCAGTGTCTGCGAG 900
Db 1095 ATGACATCTGGACCAACAGAGCTGGATGAGAGACTGTCAGTGTCTGCGAG 1154
Qy 901 CGGGCTTCGGCTGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCTATGCCAGT 960
Db 1155 CGGGCTTCGGCTGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCTATGCCAGT 1214
Qy 961 GTGTCTGTAAAAAACAACCTTTCCTCCAGCCCAATGTGGGGCCACCGAGATTTGATGAAA 1020
Db 1215 GTGTCTGTAAAAAACAACCTTTCCTCCAGCCCAATGTGGGGCCACCGAGATTTGATGAAA 1274
Qy 1021 ACACATGCGAGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCCTGGAA 1080
Db 1275 ACACATGCGAGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCCTGGAA 1334
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATCTTTGTTAAAGAAAGAAAGTTCC 1140
Db 1335 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATCTTTGTTAAAGAAAGAAAGTTCC 1394
Qy 1141 ACCACCAACATGAGCTGTACAGAGCGCATGTACAGCCCGCAGAGCTGTGAGC 1200
Db 1395 ACCACCAACATGAGCTGTACAGAGCGCATGTACAGCCCGCAGAGCTGTGAGC 1454
Qy 1201 CAGGATTTTCATATAGTGAAGAAGTGTGTGTCCTTTCATATTTGCAAGAACCCAC 1260
Db 1455 CAGGATTTTCATATAGTGAAGAAGTGTGTGTCCTTTCATATTTGCAAGAACCCAC 1514
Qy 1261 AATGAGCTAAGTGTACTGTTTTCAGTTCATCGATTTTCTATTTAGGAAACTGTGT 1320
Db 1515 AATGAGCTAAGTGTACTGTTTTCAGTTCATCGATTTTCTATTTAGGAAACTGTGT 1574
Qy 1321 TGCCACAGTAAAGTGTGTGAACAGAGAGCCCTTGTGGTCCATGCTPAACAAAGACA 1380
Db 1575 TGCCACAGTAAAGTGTGTGAACAGAGAGCCCTTGTGGTCCATGCTPAACAAAGACA 1634
Qy 1381 AAGTCTGTCTTCTGAAACATGTGGATAACTTTACAGAAATGGAGTGCATCTG 1440
Db 1635 AAGTCTGTCTTCTGAAACATGTGGATAACTTTACAGAAATGGAGTGCATCTG 1694
Qy 1441 CAAAGGCCCTCTGTAAAGACTGTTTTCGCAATGACCAACAGCCAGATTTTCCTC 1500
Db 1695 CAAAGGCCCTCTGTAAAGACTGTTTTCGCAATGACCAACAGCCAGATTTTCCTC 1754
Qy 1501 TTGTGATTTCTTTAAAAAGATGACTATATAATTTATTTCCATTAATAATTTGTTCTGC 1560
Db 1755 TTGTGATTTCTTTAAAAAGATGACTATATAATTTATTTCCATTAATAATTTGTTCTGC 1814
Qy 1561 ATTCAATTTTATACCAACAATTTGTAACCTCACTGTGATCAATATTTTATATCAT 1620
Db 1815 ATTCAATTTTATACCAACAATTTGTAACCTCACTGTGATCAATATTTTATATCAT 1874

Qy 1621 GCAAAATATGTTTAAATGAAATGAAATGTAATTTATTAATAAAAAAAAAAAAA 1674
Db 1875 GCAAAATATGTTTAAATGAAATGAAATGTAATTTATTAATAAAAAAAAAAAAA 1928

RESULT 6
AR231261 AR231261 2031 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 1 from patent US 6451764.
DEFINITION AR231261
ACCESSION AR231261
VERSION AR231261.1 GI:27272182
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2031)
AUTHORS Lee, J. and Wood, W.
TITLE VEGF-related protein
JOURNAL Patent: US 6451764-A 1 17-SEP-2002;
FEATURES Location/Qualifiers
 source
 1..2031
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
Query Match 99.4%; Score 1664.6; DB 6; Length 2031;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GTCCTTCCACCATGACATCGCTGGGCTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGGTG 60
Db 361 GTCCTTCCACCATGACATCGCTGGGCTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGGTG 420
Qy 61 CGCTGCTCCGGTCTCTCGAGAGCGCCCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Db 421 CGCTGCTCCGGTCTCTCGAGAGCGCCCGCGCGCGCTTCGAGTCCGAGCTCG 480
Qy 121 ACCTCTCGAGCGCGAGCGCCGAGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
Db 481 ACCTCTCGAGCGCGAGCGCCGAGCGCGCGCGCTTATGCAAGCAAGATCTGG 540
Qy 181 AGGACAGTTAGTGTCTGTCTCAGTGTAGATGAATCATGACTGTACTTACCCAGAT 240
Db 541 AGGACAGTTAGTGTCTGTCTCAGTGTAGATGAATCATGACTGTACTTACCCAGAT 600
Qy 241 ATTGCAATGTACAGTGTGAGTAAAGAGGCTGCGCAATACAGAGAACAG 300
Db 601 ATTGCAATGTACAGTGTGAGTAAAGAGGCTGCGCAATACAGAGAACAG 660
Qy 301 CCAACCTCAACTCAGGACAGAGAGACTATAAAATTTGTCGAGCACTTATATACAG 360
Db 661 CCAACCTCAACTCAGGACAGAGAGACTATAAAATTTGTCGAGCACTTATATACAG 720
Qy 361 AGATCTTGAAGAATTTGATTAATGAGTGGAGAAAGACTCAATGATGCCAGGAGGTG 420
Db 721 AGATCTTGAAGAATTTGATTAATGAGTGGAGAAAGACTCAATGATGCCAGGAGGTG 780
Qy 421 GTATAGATGTGGGAAGGAGTTTGGAGTCCGCAACACACCTCTTTAAACCTCATGTG 480
Db 781 GTATAGATGTGGGAAGGAGTTTGGAGTCCGCAACACACCTCTTTAAACCTCATGTG 840
Qy 481 TGTCCGTCTACAGATGTGGGGTGTGTCGAATAGTGGGGCTGAGTGCATGAACACCA 540
Db 841 TGTCCGTCTACAGATGTGGGGTGTGTCGAATAGTGGGGCTGAGTGCATGAACACCA 900
Qy 541 GCACGAGTACTCTCAGCAAGACGTTTATTTGAAATACAGTCCCTCTCTCAAGGCCCA 600
Db 901 GCACGAGTACTCTCAGCAAGACGTTTATTTGAAATACAGTCCCTCTCTCAAGGCCCA 960
Qy 601 AACAGTAAATCAGTTTTCCTCAATACACTCTCTGCGATGATGCTAACTGGATG 660
Db 961 AACAGTAAATCAGTTTTCCTCAATACACTCTCTGCGATGATGCTAACTGGATG 1020

661 TTTACAGCAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGC 720
1021 TTTACAGCAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGC 1080
721 AGGCAGGCAACAGACCTGCCACCAATACATGTGGAATAATCAATCTGCAGATGCC 780
1081 AGGCAGGCAACAGACCTGCCACCAATACATGTGGAATAATCAATCTGCAGATGCC 1140
781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGATCAACAGATGGATTCC 840
1141 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGATCAACAGATGGATTCC 1200
841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTCAAGTGTCTGCAGAG 900
1201 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTCAAGTGTCTGCAGAG 1260
901 CGGGCTTCGGCTGCCAGCTGTGGACCCACAAAGAACTAGACAGAACTCATGCCAGT 960
1261 CGGGCTTCGGCTGCCAGCTGTGGACCCACAAAGAACTAGACAGAACTCATGCCAGT 1320
961 GTGTCTGTAAAAACAACACTCTTCCCAAGCCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
1321 GTGTCTGTAAAAACAACACTCTTCCCAAGCCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1380
1021 ACATATGCCAGTGTGTATGTAAAGAACTGTGCCCCAGAAATCAACCCCTAAATCTCTGGAA 1080
1381 ACATATGCCAGTGTGTATGTAAAGAACTGTGCCCCAGAAATCAACCCCTAAATCTCTGGAA 1440
1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGATTCC 1140
1441 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGATTCC 1500
1141 ACCAACAAATGCAGCTGTACAGAGCGCCATGTACGAACCGCCAGAGAGGTTGTGAGC 1200
1501 ACCAACAAATGCAGCTGTACAGAGCGCCATGTACGAACCGCCAGAGAGGTTGTGAGC 1560
1201 CAGGATTTTCATATAGTAAGAGATGTGTCGTTGTCCTGCTCATATGGGAAAGACAC 1260
1561 CAGGATTTTCATATAGTAAGAGATGTGTCGTTGTCCTGCTCATATGGGAAAGACAC 1620
1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATATGGAATACTGTGT 1320
1621 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATATGGAATACTGTGT 1680
1321 TGCCACAGTAGAATCTGTCTGTGAACAGAGAGACCCCTGTGGGTCCATGCTAAACAGACA 1380
1681 TGCCACAGTAGAATCTGTCTGTGAACAGAGAGACCCCTGTGGGTCCATGCTAAACAGACA 1740
1381 AAAGTCTGTCTTCTGGAACCATGTGGATAAATTTACAGAAATGGACTGGAGCTCATCTG 1440
1741 AAAGTCTGTCTTCTGGAACCATGTGGATAAATTTACAGAAATGGACTGGAGCTCATCTG 1800
1441 CAAAAGCCCTCTGTGAAGCTGTTTCTGCCAATGACCAACAGCCAGAGATTTTCTCTC 1500
1801 CAAAAGCCCTCTGTGAAGCTGTTTCTGCCAATGACCAACAGCCAGAGATTTTCTCTC 1860
1501 TTGTGATTTCTTTAAAGATGACTATATAATTTATTTCCACTTAAATATTTGTTTCTGC 1560
1861 TTGTGATTTCTTTAAAGATGACTATATAATTTATTTCCACTTAAATATTTGTTTCTGC 1920
1561 ATTCAATTTTATAGCAACCAATTTGCTTAAACTCACTGTGATCAATATTTTATATCAT 1620
1921 ATTCAATTTTATAGCAACCAATTTGCTTAAACTCACTGTGATCAATATTTTATATCAT 1980
1621 GCATAATATGTTTAAATATAAATGAAATTTGTTTATTTAAATATAAATATAA 1671
1981 GCATAATATGTTTAAATATAAATGAAATTTGTTTATTTAAATATAAATATAA 2031

RESULT 7
AR342826
LOCUS 2031 bp DNA linear PAT 17-AUG-2003

DEFINITION Sequence 1 from patent US 6576608.
ACCESSION AR342826
VERSION AR342826.1 GI:33738052
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2031)
AUTHORS Lee, J. and Wood, W.
TITLE Methods of using VEGF-related protein
JOURNAL Patent: US 6576608-A 1 10-JUN-2003;
FEATURES Location/Qualifiers
1..2031
source /organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.4%; Score 1664.6; DB 6; Length 2031;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTCCTTCCACCATGCACCTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCCCGCTG 60
Db 361 GTCCTTCCACCATGCACCTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCCCGCTG 420
QY 61 CGCTGCTCCCGGTCCTCGGAGCGCGCGCGCGCGCGCTTCGAGTCCCGACTCG 120
Db 421 CGCTGCTCCCGGTCCTCGGAGCGCGCGCGCGCGCGCTTCGAGTCCCGACTCG 480
QY 121 ACCTCTGGAGCGCGAGCGCGCGCGCGCGCGCGCTTATGGAAGCAAGATCTGG 180
Db 481 ACCTCTGGAGCGCGAGCGCGCGCGCGCGCGCGCTTATGGAAGCAAGATCTGG 540
QY 181 AGGAGCAGTGTACGCTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTTACCCAGAT 240
Db 541 AGGAGCAGTGTACGCTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTTACCCAGAT 600
QY 241 ATTGGAATAATGTAAGATGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAG 300
Db 601 ATTGGAATAATGTAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAG 660
QY 301 CCAACCTCAACTCAAGACAGAGAGACTATAAATTTGCTGAGACATTAATAACAG 360
Db 661 CCAACCTCAACTCAAGACAGAGAGACTATAAATTTGCTGAGACATTAATAACAG 720
QY 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 420
Db 721 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 780
QY 421 GTATAGATGTGGGAGAGAGTGTGGAGTCCGCAAAACACCTCTTTTAAACCTCCATGTG 480
Db 781 GTATAGATGTGGGAGAGAGTGTGGAGTCCGCAAAACACCTCTTTTAAACCTCCATGTG 840
QY 481 TGTCCGTCTACAGATGTGGGGTGTGCTCAATAGTGGGGCTGCAAGTGCATGAACACCA 540
Db 841 TGTCCGTCTACAGATGTGGGGTGTGCTCAATAGTGGGGCTGCAAGTGCATGAACACCA 900
QY 541 GCACAGCTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGGCCCCA 600
Db 901 GCACAGCTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGGCCCCA 960
QY 601 AACCAGTAAACAATCAGTTTTTGCACATCACAATCTTCCTGCGATGCATGTCTTAACTGGATG 660
Db 961 AACCAGTAAACAATCAGTTTTTGCACATCACAATCTTCCTGCGATGCATGTCTTAACTGGATG 1020
QY 661 TTTACAGCAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGC 720
Db 1021 TTTACAGCAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGC 1080
QY 721 AGGCAGGCAACAGACCTGCCACCAATTTACATGTGGAATAATCAATCTGCAGATGCC 780
Db 1081 AGGCAGGCAACAGACCTGCCACCAATTTACATGTGGAATAATCAATCTGCAGATGCC 1140

```
QY 781 TGGCTCAGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATCC 840
Db 1141 TGGCTCAGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATCC 1200
QY 841 ATGACATCTGTGGACCAACCAAGAGCTGGATGAAGAGACCTGTCAGTGTCTGCGAGAG 900
Db 1201 ATGACATCTGTGGACCAACCAAGAGCTGGATGAAGAGACCTGTCAGTGTCTGCGAGAG 1260
QY 901 CGGGCTTTGGCCTGCCAGCTGTGGAGCCCAACCAAGAACTAGACAGAAATCATGCCAGT 960
Db 1261 CGGGCTTTGGCCTGCCAGCTGTGGAGCCCAACCAAGAACTAGACAGAAATCATGCCAGT 1320
QY 961 GTGCTGTAAACCAACCTCTTCCCGCCCAATGTCGGGCCCAACCGGAATTTGATGAAA 1020
Db 1321 GTGCTGTAAACCAACCTCTTCCCGCCCAATGTCGGGCCCAACCGGAATTTGATGAAA 1380
QY 1021 ACACATGCCAGTGTGTATGTAAGAGACCTGCCGCCAGAAATCAACCCCTAAATCCTGGAA 1080
Db 1381 ACACATGCCAGTGTGTATGTAAGAGACCTGCCGCCAGAAATCAACCCCTAAATCCTGGAA 1440
QY 1081 AATGTCCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
Db 1441 AATGTCCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAGGAAAGAGTTCC 1500
QY 1141 ACCACCAACATGAGCTGTGTACAGAGCGGCATGTACGACCGCCAGAGCTTTGTGAGC 1200
Db 1501 ACCACCAACATGAGCTGTGTACAGAGCGGCATGTACGACCGCCAGAGCTTTGTGAGC 1560
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTCCTTCATATTTGGCAAGACCCAC 1260
Db 1561 CAGGATTTTCATATAGTGAAGAGTGTGCTGTCCTTCATATTTGGCAAGACCCAC 1620
QY 1261 AAATGAGCTAAGATTTACTGTTTCCAGTTCATCGATTTCTTATTTGAAAGACTGTGT 1320
Db 1621 AAATGAGCTAAGATTTACTGTTTCCAGTTCATCGATTTCTTATTTGAAAGACTGTGT 1680
QY 1321 TGCCACAGTGAAGCTGTGTGAACAGAGAGACCTTTGGTGCCATGCTTAACAAAGACA 1380
Db 1681 TGCCACAGTGAAGCTGTGTGAACAGAGAGACCTTTGGTGCCATGCTTAACAAAGACA 1740
QY 1381 AAGTCTGCTTCTTCCAGCATGCTGTGAACATGCTTAACATTTCTTATTTGAAAGACTGTGT 1440
Db 1741 AAGTCTGCTTCTTCCAGCATGCTGTGAACATGCTTAACATTTCTTATTTGAAAGACTGTGT 1800
QY 1441 CAAAGGCTCTTGTGAAGACTGTTTCTGCCAATGACCAACAGCCAGAGATTTTCCCTC 1500
Db 1801 CAAAGGCTCTTGTGAAGACTGTTTCTGCCAATGACCAACAGCCAGAGATTTTCCCTC 1860
QY 1501 TTGTGATTTCTTTAAAGAGATGACTATATATTTTATTTTCCACTAAATAATTTCTTCTGC 1560
Db 1861 TTGTGATTTCTTTAAAGAGATGACTATATATTTTATTTTCCACTAAATAATTTCTTCTGC 1920
QY 1561 ATTCAATTTTATAGCAACAAATTTGTAAGACTGCTGATGATCAATATTTTATATCAT 1620
Db 1921 ATTCAATTTTATAGCAACAAATTTGTAAGACTGCTGATGATCAATATTTTATATCAT 1980
QY 1621 GCATAATATGTTTAAATAAATAAGAAATTTGATTTTATAAAAAAATAAAAAA 1671
Db 1981 GCATAATATGTTTAAATAAATAAGAAATTTGATTTTATAAAAAAATAAAAAA 2031

RESULT 8
BC063685 LOCUS
DEFINITION Homo sapiens cDNA clone MGC:74703 IMAGE:4513934, complete cds.
ACCESSION BC063685
VERSION BC063685.1 GI:39645812
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2076)
```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, E., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

JOURNAL

2 (bases 1 to 2076)
Strausberg, R.
Direct Submission

PUBMED

Submitted (08-DEC-2003) National Institutes of Health, Mammalian

REFERENCE

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov

REMARK

Tissue Procurement: ATCC

COMMENT

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Achter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaskey, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantrispop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 134 Row: j Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19924300.

FEATURES

source

1..2076
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:74703 IMAGE:4513934"
/tissue_type="Bladder, transitional cell papilloma"
/clone_lib="NIH MGC_93"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
417..1676
/codon_start=1
/product="unknown (protein for MGC:74703)"
/protein_id="AAH63685.1"
/db_xref="GI:39645813"

CDS

RVHSIIIRSLPATLPOCOAANKTCTPTNWNHICRLAQEDFMSDAGDSTDGF
HDICGNKLEDETCOCVACRAGLRPASCGPHKELDENSCOCUCNKLPSQCGANREF
DENTCOCVKETCPNPNPGKACACECTESPOKLLKKKHHQTCSTIRPFCINRQ
KACEPFSVSEECVRCPVSWKSPQMS"
801..1055
/note="PDGF; Region: Platelet-derived and vascular
endothelial growth factors (PDGF, VEGF) family"
/db_xref="CDD:smart00141"

misc_feature

ORIGIN

Query Match 99.4%; Score 1664.6; DB 9; Length 2076;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTCTCTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGCTG 60
DB 406 GTCTCTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGCTG 465
QY 61 CGTGTCTCCGGGTCTCTCGGAGGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
DB 466 CGTGTCTCCGGGTCTCTCGGAGGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 525
QY 121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 180
DB 526 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 585
QY 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTCTACCCAGAA 240
DB 586 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTCTACCCAGAA 645
QY 241 ATTGGAATATGACAGTGTACGTAAGGAAGAGGCTGGCAACATAACAGAGAACAGG 300
DB 646 ATTGGAATATGACAGTGTACGTAAGGAAGAGGCTGGCAACATAACAGAGAACAGG 705
QY 301 CCAACCTCAACTCAAGGACAGAGACTATAAATTTGCTGCAGCAACATTAATACAG 360
DB 706 CCAACCTCAACTCAAGGACAGAGACTATAAATTTGCTGCAGCAACATTAATACAG 765
QY 361 AGATCTTGAAGATATTGATAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTGT 420
DB 766 AGATCTTGAAGATATTGATAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTGT 825
QY 421 GTATAGATGTGGGGAAGGAGTTTGGAGTCCGACAAACACCTTCTTTAACTCCATCTG 480
DB 826 GTATAGATGTGGGGAAGGAGTTTGGAGTCCGACAAACACCTTCTTTAACTCCATCTG 885
QY 481 TGTCCGTCTACAGATGTGGGGTGTGTGCAATAGTGGGGGTGAGGCTGCATGCAACCA 540
DB 886 TGTCCGTCTACAGATGTGGGGTGTGTGCAATAGTGGGGGTGAGGCTGCATGCAACCA 945
QY 541 GCACGAGCTACCTCAGCAGAGCGTTATTGAAATTTACAGTGCCTCTCTCAGGCCCA 600
DB 946 GCACGAGCTACCTCAGCAGAGCGTTATTGAAATTTACAGTGCCTCTCTCAGGCCCA 1005
QY 601 AACCAGTAACAAATCAGTTTGGCAATCACACTTCTGCGGATGCATGCTAACTGGATG 660
DB 1006 AACCAGTAACAAATCAGTTTGGCAATCACACTTCTGCGGATGCATGCTAACTGGATG 1065
QY 661 TTACAGACAGTTTATTCATTTAGAGCTTCCCTGCGAGCAACACTACACAGTGTG 720
DB 1066 TTACAGACAGTTTATTCATTTAGAGCTTCCCTGCGAGCAACACTACACAGTGTG 1125
QY 721 AGGACGCAACAGACTCCGCCACCAATTACATGTGGAATAATCACATCTGCAGATGC 780
DB 1126 AGGACGCAACAGACTCCGCCACCAATTACATGTGGAATAATCACATCTGCAGATGC 1185
QY 781 TGGCTCAGGAAGATTTATTTTCTCGGATGCTGGAGTGAATCAACAGATGGAATTC 840
DB 1186 TGGCTCAGGAAGATTTATTTTCTCGGATGCTGGAGTGAATCAACAGATGGAATTC 1245
QY 841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGCTGTCTGCAGAG 900
DB 1246 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGCTGTCTGCAGAG 1305

QY 901 CGGGGCTTTGGGCTTGGCCTGCGAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
DB 1306 CGGGGCTTTGGGCTTGGCCTGCGAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 1365
QY 961 GTGCTGTAAACAACTCTTCCCGAGCGCAATGTGGGGCCCAACCGAGAAATTTTCATGAAA 1020
DB 1366 GTGCTGTAAACAACTCTTCCCGAGCGCAATGTGGGGCCCAACCGAGAAATTTTCATGAAA 1425
QY 1021 ACACATGCGAGTGTGTATGTAAAGAACTGCGCCAGAAATCAACCCCTAAATCCTCGAA 1080
DB 1426 ACACATGCGAGTGTGTATGTAAAGAACTGCGCCAGAAATCAACCCCTAAATCCTCGAA 1485
QY 1081 AATGCTGCTGTGATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTC 1140
DB 1486 AATGCTGCTGTGATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTC 1545
QY 1141 ACCACCAACATGAGCTGTTCACAGCGGCGCATGTACGAACCCGACAGAGCTTGTGAGC 1200
DB 1546 ACCACCAACATGAGCTGTTCACAGCGGCGCATGTACGAACCCGACAGAGCTTGTGAGC 1605
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATTTGGCAAGACCAC 1260
DB 1606 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATTTGGAAAGACCAC 1665
QY 1261 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1320
DB 1666 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1725
QY 1321 TGCCACAGTGAAGTGTGTGTGAACAGAGAGACCCCTTGTGGTCCATGCTTAAACAAAGACA 1380
DB 1726 TGCCACAGTGAAGTGTGTGTGAACAGAGAGACCCCTTGTGGTCCATGCTTAAACAAAGACA 1785
QY 1381 AAAGTGTCTTCTTCTGCAACCATGTGTGATAACTTTACAGAAATGGAGTGGAGTCTCTG 1440
DB 1786 AAAGTGTCTTCTTCTGCAACCATGTGTGATAACTTTACAGAAATGGAGTGGAGTCTCTG 1845
QY 1441 CAAAGGCTCTTGTAAAGACTGTTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1500
DB 1846 CAAAGGCTCTTGTAAAGACTGTTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1905
QY 1501 TTGTGATTTCTTTAAAGATGACTATATAATTTTATTTTCCACTTAAATAATTTTCTTCTG 1560
DB 1906 TTGTGATTTCTTTAAAGATGACTATATAATTTTATTTTCCACTTAAATAATTTTCTTCTG 1965
QY 1561 ATTCAATTTTATAGCAACCAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT 1620
DB 1966 ATTCAATTTTATAGCAACCAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT 2025
QY 1621 GCAAAATATGTTTAAATAATAATGAAATTTGTAATTTTATAAAAAA 1671
DB 2026 GCAAAATATGTTTAAATAATAATGAAATTTGTAATTTTATAAAAAA 2076

RESULT 9

AR231262/c

LOCUS

DEFINITION

AR231262

SEQUENCE

AR231262.1

GI:27272183

KEYWORDS

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 2031)

AUTHORS

Lee, J. and Wood, W.

TITLE

VEGF-related protein

JOURNAL

Patent: US 6451764-A 2 17-SEP-2002;

FEATURES

Location/Qualifiers

1..2031

source

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

2031 bp

DNA

Sequence 2 from patent US 6451764.

linear

PAT 20-DEC-2002

Query Match		99.3%; Score 1663; DB 6; Length 2031;
Best Local Similarity		99.7%; Pred. No. 0;
Matches 1666; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
QY	1	GTCTTCCACCATGACATCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCCGCGCTG 60
Db	1671	GTCTTCCACCATGACATCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCCGCGCTG 1612
QY	61	CGTGCTCCGGTCTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db	1611	TGCTCTCCGGTCTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1552
QY	121	ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db	1551	ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1492
QY	181	AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAT 240
Db	1491	AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAT 1432
QY	241	ATTGGAAATCTACAAGTCTCAGCTTAAGGAAGAGCGCTGGCAACATAACAGAAACAG 300
Db	1431	ATTGGAAATCTACAAGTCTCAGCTTAAGGAAGAGCGCTGGCAACATAACAGAAACAG 1372
QY	301	CCAACTCTCAACTCAAGGACAGAGACTATAAAATTTGCTGCAGACATTTATAATACAG 360
Db	1371	CCAACTCTCAACTCAAGGACAGAGACTATAAAATTTGCTGCAGACATTTATAATACAG 1312
QY	361	AGATCTTGAAAGTATTGATAATGAGTGGAAAGACTCAATGATGCCACGGAGGTGT 420
Db	1311	AGATCTTGAAAGTATTGATAATGAGTGGAAAGACTCAATGATGCCACGGAGGTGT 1252
QY	421	GTATAGATGTGGGAAGAGTTTGGAGTTCGCGACCAACACTTCTTTAAACCTCCATGTG 480
Db	1251	GTATAGATGTGGGAAGAGTTTGGAGTTCGCGACCAACACTTCTTTAAACCTCCATGTG 1192
QY	481	TGTCGCTCAGATGTGGGGTTGCTGCAATAGTAGGGGCTGCAGTGCATGAACACCA 540
Db	1191	TGTCGCTCAGATGTGGGGTTGCTGCAATAGTAGGGGCTGCAGTGCATGAACACCA 1132
QY	541	GCAGAGCTACCTCAGCAAGAGCTTATTGAAATACAGTGCCTCTCTCAGAGGCCCA 600
Db	1131	GCAGAGCTACCTCAGCAAGAGCTTATTGAAATACAGTGCCTCTCTCAGAGGCCCA 1072
QY	601	AACAGTAAACATCAGTTTTGCCAATCACACTTCTCTGCCGATGCATGTCTAACTGGATG 660
Db	1071	AACAGTAAACATCAGTTTTGCCAATCACACTTCTCTGCCGATGCATGTCTAACTGGATG 1012
QY	661	TTTACAGACAGTTCAATTCATTTAGAGCTTCCCTGCCAGCAACACTACCAAGTGT 720
Db	1011	TTTACAGACAGTTCAATTCATTTAGAGCTTCCCTGCCAGCAACACTACCAAGTGT 952
QY	721	AGGAGCGAAACAAGACTCGCCCAACCAATACATGTGGAAATATCACATCTGCAGATGC 780
Db	951	AGGAGCGAAACAAGACTCGCCCAACCAATACATGTGGAAATATCACATCTGCAGATGC 992
QY	781	TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGTGAATCAACAGATGGAATCC 840
Db	891	TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGTGAATCAACAGATGGAATCC 832
QY	841	ATGACATCTGTGGACCAACAAGAGCTGTGATGAAGAGCTGTGATGTCTGCAGAG 900
Db	831	ATGACATCTGTGGACCAACAAGAGCTGTGATGAAGAGCTGTGATGTCTGCAGAG 772
QY	901	CGGGCTTCGGCTGCGCAGCTGGAGCCCAACAAAGAACTAGACAAACTCATGCGAGT 960
Db	771	CGGGCTTCGGCTGCGCAGCTGGAGCCCAACAAAGAACTAGACAAACTCATGCGAGT 712
QY	961	GTGTCTGTAACAAACTCTTCCCGACGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
Db	711	GTGTCTGTAACAAACTCTTCCCGACGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 652
QY	1021	ACACATGCCAGTGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTCTGGAA 1080
Db	651	ACACATGCCAGTGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTCTGGAA 592
QY	1081	AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
Db	591	AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 532
QY	1141	ACCACCAACATGACGCTGTTACAGACGCGCATGTACGAACCGCCAGAGGCTTTGTGAGC 1200
Db	531	ACCACCAACATGACGCTGTTACAGACGCGCATGTACGAACCGCCAGAGGCTTTGTGAGC 472
QY	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATGGGAAAGCCAC 1260
Db	471	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATGGGAAAGCCAC 412
QY	1261	AAATGAGCTTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATATGGAATACTGTGT 1320
Db	411	AAATGAGCTTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATATGGAATACTGTGT 352
QY	1321	TGCCACAGTAGAATCTGTCTGTGAACAGAGAGACCTTGTGGTCCATGCTAACAAAGACA 1380
Db	351	TGCCACAGTAGAATCTGTCTGTGAACAGAGAGACCTTGTGGTCCATGCTAACAAAGACA 292
QY	1381	AAAGTCTGTCTTCTCTGAACCATGTGGATACTTTACAGAAATGACTGGAGCTCATCTG 1440
Db	291	AAAGTCTGTCTTCTCTGAACCATGTGGATACTTTACAGAAATGACTGGAGCTCATCTG 232
QY	1441	CAAAAGGCTCTGTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1500
Db	231	CAAAAGGCTCTGTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 172
QY	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Db	171	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 112
QY	1561	ATTCAATTTTATAGCAACAACATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Db	111	ATTCAATTTTATAGCAACAACATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 52
QY	1621	GCAAAATATGTTTAAATTAATAATGAANAATGATTTATTAATAAAAAA 1671
Db	51	GCAAAATATGTTTAAATTAATAATGAANAATGATTTATTAATAAAAAA 1
RESULT 10		
AR342827/c		
LOCUS		AR342827 2031 bp DNA linear PAT 17-AUG-2003
DEFINITION		Sequence 2 from patent US 6576608.
ACCESSION		AR342827
VERSION		AR342827.1 GI:33738053
KEYWORDS		
SOURCE		Unknown.
ORGANISM		Unknown.
REFERENCE		Unclassified.
AUTHORS		1 (bases 1 to 2031)
TITLE		Lee, J. and Wood, W.
JOURNAL		Methods of using VEGF-related protein
FEATURES		Patent: US 6576608-A 2 10-JUN-2003;
source		Location/Qualifiers
ORIGIN		1..2031
Query Match		/organism="unknown"
Best Local Similarity		/mol_type="genomic DNA"
Matches 1666; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
QY	1	GTCTTCCACCATGACATCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCCGCGCTG 60
Db	1671	GTCTTCCACCATGACATCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCCGCGCTG 1612


```
QY 61 CGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCTTCGAGTCGGAGCTCG 120
Db 1611 TGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCTTCGAGTCGGAGCTCG 1552
QY 121 ACCTCTCGAGCGGGAGCGCGAGCGGGAGCGCGCGCTTATGCAAGCAAGATCTCG 180
Db 1551 ACCTCTCGAGCGGGAGCGCGAGCGGGAGCGCGCGCTTATGCAAGCAAGATCTCG 1492
QY 181 AGAGCAGTTACGGTCTGTCTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 240
Db 1491 AGAGCAGTTACGGTCTGTCTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 1432
QY 241 ATTGGAATATGTAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAG 300
Db 1431 ATTGGAATATGTAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAG 1372
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGTCGAGCACAATTAATACAG 360
Db 1371 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGTCGAGCACAATTAATACAG 1312
QY 361 AGATCTGAAAGTATTGATTAAGTGTGGAAGAGTCAATGCTGCCACGGAGGTGT 420
Db 1311 AGATCTGAAAGTATTGATTAAGTGTGGAAGAGTCAATGCTGCCACGGAGGTGT 1252
QY 421 GTATAGATGTGGGAGAGGAGTTTGGAGTCGCGACAAACACCTTCTTTAAACCTCCATGTG 480
Db 1251 GTATAGATGTGGGAGAGGAGTTTGGAGTCGCGACAAACACCTTCTTTAAACCTCCATGTG 1192
QY 481 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTGAAGGCTGCAAGTCAATGACCA 540
Db 1191 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTGAAGGCTGCAAGTCAATGACCA 1132
QY 541 GCACGAGTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Db 1131 GCACGAGTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 1072
QY 601 AACGAGTAAACATGAGTTTGGCAATCACAATCTCTGCGGATGCAATGCTAACTGGAGT 660
Db 1071 AACGAGTAAACATGAGTTTGGCAATCACAATCTCTGCGGATGCAATGCTAACTGGAGT 1012
QY 661 TTTACAGACAAGTTTATCTCATTTAGAGTGTCCCTGCCAGCAACACTACCACAGTGT 720
Db 1011 TTTACAGACAAGTTTATCTCATTTAGAGTGTCCCTGCCAGCAACACTACCACAGTGT 952
QY 721 AGGACGGAACAAGACTGCCCAACCAATTAATGGAATATCAATCATCTCAGATGCC 780
Db 951 AGGACGGAACAAGACTGCCCAACCAATTAATGGAATATCAATCATCTCAGATGCC 892
QY 781 TGGCTCAGGAAGATTATTTATCTTCTCGATGCTGGAGTGAATCAATCAATGAGTTC 840
Db 891 TGGCTCAGGAAGATTATTTATCTTCTCGATGCTGGAGTGAATCAATCAATGAGTTC 832
QY 841 ATGACATCTGTGGACCAACAAGAGAGTGGATGAAGAGACTCTGCTGAGTGTCTGCAGAG 900
Db 831 ATGACATCTGTGGACCAACAAGAGAGTGGATGAAGAGACTCTGCTGAGTGTCTGCAGAG 772
QY 901 CGGGCTTCGGCTCGCCAGCTGTGGACCCCAAGAACTAGACAGAAATCTATGCCAGT 960
Db 771 CGGGCTTCGGCTCGCCAGCTGTGGACCCCAAGAACTAGACAGAAATCTATGCCAGT 712
QY 961 GTGTCTGTAATAAACAATCTCTCCAGCAATGTGGGGCAACCGAGAAATTTGATGAA 1020
Db 711 GTGTCTGTAATAAACAATCTCTCCAGCAATGTGGGGCAACCGAGAAATTTGATGAA 552
QY 1021 ACATGCCAGTGTATGTAAGAACTCTGCCCAAGAAATCAACCCCTAAATCTCGGAA 1080
Db 651 ACATGCCAGTGTATGTAAGAACTCTGCCCAAGAAATCAACCCCTAAATCTCGGAA 592
QY 1081 AATGTCTGTGAATGTACAGAACTCCAGAAATGCTTTGTAAGGAAGAAAGTCTCC 1140
Db 591 AATGTCTGTGAATGTACAGAACTCCAGAAATGCTTTGTAAGGAAGAAAGTCTCC 532
QY 1141 ACCACCAACATGAGTGTATACAGAGCGGCTATGACGACCGCCAGAGGCTTGTGAGC 1200
```

```
Db 531 ACCACCAACATGAGTGTATACAGAGCGGCTATGACAAACCGCAGAGGCTTGTGAGC 472
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 471 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 412
QY 1261 AAATGAGCTAAGATGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 411 AAATGAGCTAAGATGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 352
QY 1321 TGCCACAGTAGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 351 TGCCACAGTAGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 292
QY 1381 AAAGTGTGTCTTCTCTGAAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 291 AAAGTGTGTCTTCTCTGAAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 232
QY 1441 CAAAAGGCTCTTCTGTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 231 CAAAAGGCTCTTCTGTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 172
QY 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTTATTTCCACTAAATAATTTGTTCCTGC 1560
Db 171 TTGTGATTTCTTTAAAGAAATGACTATATAATTTTATTTCCACTAAATAATTTGTTCCTGC 112
QY 1561 ATTCAATTTTATAGCAACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 111 ATTCAATTTTATAGCAACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 52
QY 1621 GCAAAATATGTTTAAATATAAAATGAAATTTGATTTTATAAAAAA 1671
Db 51 GCAAAATATGTTTAAATATAAAATGAAATTTGATTTTATAAAAAA 1

RESULT 11
AR112510
LOCUS AR112510 1997 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 7 from patent US 6130071.
ACCESSION AR112510
VERSION AR112510.1 GI:14092410
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Alitalo, K. and Joukov, V.
TITLE Vascular endothelial growth factor C (VEGF-C) .DELTA. Cys.sub.156
JOURNAL Protein and gene, and uses thereof
FEATURES
    Location/Qualifiers
        source
            1..1997
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCTCTCCACCATGACACTCGCTGGGCTTCTTCTGTGGGGTGTCTGTCTGCCCTG 60
Db 341 GTCTCTCCACCATGACACTCGCTGGGCTTCTTCTGTGGGGTGTCTGTCTGCCCTG 400
QY 61 CGCTGTCTCCCGGTCCTCGAGAGCGCCCGCGCGCGCTTCGAGTCCGAGATCG 120
Db 401 CGCTGTCTCCCGGTCCTCGAGAGCGCCCGCGCGCGCTTCGAGTCCGAGATCG 460
QY 121 ACCCTCGAGAGCGGAGCCCGAGCGGGCGAGGCGCGCTTATGCAAGCAAGATCTGG 180
Db 461 ACCCTCGAGAGCGGAGCCCGAGCGGGCGAGGCGCGCTTATGCAAGCAAGATCTGG 520
```



```
QY 181 AGGACAGTACGGTCTGTGTCCAGTGTAGATGAATCATCACTGTACTTACCCAGAT 240
Db 521 AGGACAGTACGGTCTGTGTCCAGTGTAGATGAATCATCACTGTACTTACCCAGAT 580
QY 241 ATTGAAAATGTACAAAGTGTAGCTTAAGAAAGAGGCTGCAACATACAGAGAAACAGG 300
Db 581 ATTGAAAATGTACAAAGTGTAGCTTAAGAAAGAGGCTGCAACATACAGAGAAACAGG 640
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACTTATAATACAG 360
Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACTTATAATACAG 700
QY 361 AGATCTTCAAGATTTGATATGATGCGAGAGAGACTCAATGCTGCGAGGAGTGT 420
Db 701 AGATCTTCAAGATTTGATATGATGCGAGAGAGACTCAATGCTGCGAGGAGTGT 760
QY 421 GTATAGATGTGGGAGAGGTTGGAGTTCGGGACAAACACACCTTCTTAAACCTCCATGTG 480
Db 761 GTATAGATGTGGGAGAGGTTGGAGTTCGGGACAAACACACCTTCTTAAACCTCCATGTG 820
QY 481 TGTCCGTCTACAGATGTGGGAGTGTGCTGCAATAGTAGGGGCTGAGTGCATGAACACCA 540
Db 821 TGTCCGTCTACAGATGTGGGAGTGTGCTGCAATAGTAGGGGCTGAGTGCATGAACACCA 880
QY 541 GCACGAGTACTCTCAGCAAGAGCTTATTTGAAATTTACAGTCCCTCTCTCTCAAGGCCCA 600
Db 881 GCACGAGTACTCTCAGCAAGAGCTTATTTGAAATTTACAGTCCCTCTCTCTCAAGGCCCA 940
QY 601 AACCAATCAATCAATTTTCCCAATCACTCTCTGCGGATGATGTCTAAACCTGGATG 660
Db 941 AACCAATCAATCAATTTTCCCAATCACTCTCTGCGGATGATGTCTAAACCTGGATG 1000
QY 661 TTTACAGCAAGTTCATTCATTTATAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720
Db 1001 TTTACAGCAAGTTCATTCATTTATAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 1060
QY 721 AGGCAGCAACAGCACTGCCCAATACATTTGAAATTTACAGTCCCTCTCTCTCAAGGCCCA 780
Db 1061 AGGCAGCAACAGCACTGCCCAATACATTTGAAATTTACAGTCCCTCTCTCTCAAGGCCCA 1120
QY 781 TGGCTCAGGAGATTTATGTTTCTCGGATGCTGAGATGACTCAACAGATGATTC 840
Db 1121 TGGCTCAGGAGATTTATGTTTCTCGGATGCTGAGATGACTCAACAGATGATTC 1180
QY 841 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGACAGAG 900
Db 1181 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGACAGAG 1240
QY 901 CGGGCTTCGGCTGCCAGTGTGGACCCCAACAAAGAACTAGACAGAAACTCATGCGAGT 960
Db 1241 CGGGCTTCGGCTGCCAGTGTGGACCCCAACAAAGAACTAGACAGAAACTCATGCGAGT 1300
QY 961 GTGTCTGTAAAAACAACTCTTCCCGCCCAATGTGGGCGCAACCGAGAAATTTGATGAA 1020
Db 1301 GTGTCTGTAAAAACAACTCTTCCCGCCCAATGTGGGCGCAACCGAGAAATTTGATGAA 1360
QY 1021 ACACATGCCAGTGTGTATGTAAAAAGAACTTCCCGCCCAATGTAAAACTTAACTCTGAA 1080
Db 1361 ACACATGCCAGTGTGTATGTAAAAAGAACTTCCCGCCCAATGTAAAACTTAACTCTGAA 1420
QY 1081 AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGTTC 1140
Db 1421 AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGTTC 1480
QY 1141 ACCCAACATGAGCTGTTTACAGAGGCGCATGTAGCAACCGCAGAGGCTTGTGAGC 1200
Db 1481 ACCCAACATGAGCTGTTTACAGAGGCGCATGTAGCAACCGCAGAGGCTTGTGAGC 1540
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTGCTTCAATATGCAAGAACAC 1260
Db 1541 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTGCTTCAATATGCAAGAACAC 1600
QY 1261 AATGAGCTAAGATTTGTACTCTTTTCCAGTTTCACTCGATTTTCTATTATGAAACCTGTGT 1320
```

```
Db 1601 AAATGAGCTAAGATTTACTGTGTTTCCAGTTCATCGATTTCTATTATGAAAACTGTGT 1660
QY 1321 TGCACAGTAGAAGTCTGTGTGAACAGAGAGACCTTGTGGTCCATGTCTAACAAAGACA 1380
Db 1661 TGCACAGTAGAAGTCTGTGTGAACAGAGAGACCTTGTGGTCCATGTCTAACAAAGACA 1720
QY 1381 AAAGTCTGTCTTTTCCGAACCACTGTGGATAAATTTTACAGAAATGGAGTCACTGTG 1440
Db 1721 AAAGTCTGTCTTTTCCGAACCACTGTGGATAAATTTTACAGAAATGGAGTCACTGTG 1780
QY 1441 CAAAAGCCCTCTGTGAAGACTGTTTTCTGCCAATGACCAACAGCCCAAGATTTTCCCTC 1500
Db 1781 CAAAAGCCCTCTGTGAAGACTGTTTTCTGCCAATGACCAACAGCCCAAGATTTTCCCTC 1840
QY 1501 TTGTGATTTCTTTTAAAGAAATGACTATATAATTTTCCACTTAAATAATTTTCTGTCG 1560
Db 1841 TTGTGATTTCTTTTAAAGAAATGACTATATAATTTTCCACTTAAATAATTTTCTGTCG 1900
QY 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCACTGTGATCAATTTTATATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACAAATTTGGTAAACTCACTGTGATCAATTTTATATATCAT 1960
QY 1621 GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1655
Db 1961 GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1995
```

RESULT 12

```
AR157589
LOCUS AR157589 1997 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 44 from patent US 6245530.
ACCESSION AR157589
VERSION AR157589.1 GI:16218541
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Alitalo, K. and Joukov, V.
TITLE Receptor ligand
JOURNAL Patent: US 6245530-A 44 12-JUN-2001;
FEATURES
    source
        1..1997
            /organism="unknown"
            /mol_type="unassigned DNA"
```

ORIGIN

```
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCCCTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGGTG 60
Db 341 GTCCCTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGGTG 400
QY 61 CGCTGCTCCCGGGTCTCTCGCAGGCGCCCGCGCGCGCGCGCTTCGAGTCCGAGACTCG 120
Db 401 CGCTGCTCCCGGGTCTCTCGCAGGCGCCCGCGCGCGCGCTTCGAGTCCGAGACTCG 460
QY 121 ACCTCTCGACCGGAGCGCCCGCGCGCGCGCGCGCTTCGAGTCCGAGACTCG 180
Db 461 ACCTCTCGACCGGAGCGCCCGCGCGCGCGCGCGCTTCGAGTCCGAGACTCG 520
QY 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTCTTACCAGAAAT 240
Db 521 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTCTTACCAGAAAT 580
QY 241 ATTGAAAATGTACAAAGTGTAGCTTAAGAAAGAGGCTGCAACATACAGAGAAACAGG 300
Db 581 ATTGAAAATGTACAAAGTGTAGCTTAAGAAAGAGGCTGCAACATACAGAGAAACAGG 640
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACTTATAATACAG 360
```

Db 641 CCAACCTCACTCAAGGACAGAGAGACTATAAATTTGCTGACGACATTAATAACAG 700
Qy 361 AGATCTTTGAAAGATTGATAATAGTGGAGAAAGACTCAATGCATGCCACGGGAGTGT 420
Db 701 AGATCTTTGAAAGATTGATAATAGTGGAGAAAGACTCAATGCATGCCACGGGAGTGT 760
Qy 421 GTATAGATGGGGAAGAGTTGGAGTCGGGACAAACACCTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGGGGAAGAGTTGGAGTCGGGACAAACACCTCTTTAAACCTCCATGTG 820
Qy 481 TGTCCGTCTACAGATGTGGGGTTGCTGCAATAGTGGGGGCTGCAGTGCATGAACACCA 540
Db 821 TGTCCGTCTACAGATGTGGGGTTGCTGCAATAGTGGGGGCTGCAGTGCATGAACACCA 880
Qy 541 GCAGAGTACCTCAGCAGACGTTATTTGAATACAGTGCCTCTCTCTCAAGGCCCA 600
Db 881 GCAGAGTACCTCAGCAGACGTTATTTGAATACAGTGCCTCTCTCTCAAGGCCCA 940
Qy 601 AACCAATAACATCAGTTTGGCAATCACCTTCTCCGATGCATGTCTAAACTGGATG 660
Db 941 AACCAATAACATCAGTTTGGCAATCACCTTCTCCGATGCATGTCTAAACTGGATG 1000
Qy 661 TTTACAGAAAGTTTCATTCATTAATAGAGTTCCTCCGAGCAAACTACCAACAGTGC 720
Db 1001 TTTACAGAAAGTTTCATTCATTAATAGAGTTCCTCCGAGCAAACTACCAACAGTGC 1060
Qy 721 AGGAGGGAACAAGACCTGCCCAACCAATTACATGTGGAATATCATCTGCAGATGCC 780
Db 1061 AGGAGGGAACAAGACCTGCCCAACCAATTACATGTGGAATATCATCTGCAGATGCC 1120
Qy 781 TGGCTCAGGAAGATTTATGTTTCTCCGATGTGGAGTGAATCAACAGATGGATTC 840
Db 1121 TGGCTCAGGAAGATTTATGTTTCTCCGATGTGGAGTGAATCAACAGATGGATTC 1180
Qy 841 ATGATCTGTGGACCAACAAGAGTGTGATGAGAGACCTGTCAGTGTCTGCAGAG 900
Db 1181 ATGATCTGTGGACCAACAAGAGTGTGATGAGAGACCTGTCAGTGTCTGCAGAG 1240
Qy 901 CGGGCTTCGGCTGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1241 CGGGCTTCGGCTGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 1300
Qy 961 GTGTCTGTAAAAACAACCTCTTCCAGCCAAATGTGGGGCCAAACCGAGAATTTGATGAAA 1020
Db 1301 GTGTCTGTAAAAACAACCTCTTCCAGCCAAATGTGGGGCCAAACCGAGAATTTGATGAAA 1360
Qy 1021 ACATATGCCAGTGTATGTAAAGAGACCTGCCCAAGAACTCAACCCCTAAATCCTGGAA 1080
Db 1361 ACATATGCCAGTGTATGTAAAGAGACCTGCCCAAGAACTCAACCCCTAAATCCTGGAA 1420
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAAAGTTCC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAAAGTTCC 1480
Qy 1141 ACCACCAACAATGACAGTGTATACAGAGGCGCATGTAGAACCCGACAGAGGCTTGTGAGC 1200
Db 1481 ACCACCAACAATGACAGTGTATACAGAGGCGCATGTAGAACCCGACAGAGGCTTGTGAGC 1540
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTGTCCTTCATATTTGCAAGAACCCAC 1260
Db 1541 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTGTCCTTCATATTTGCAAGAACCCAC 1600
Qy 1261 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTTCTATTATGAAACCTGTGT 1320
Db 1601 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTTCTATTATGAAACCTGTGT 1660
Qy 1321 TGCCACAGTAGAATGTCTGTGAACAGAGAGCCCTTGTGGGTCATGTCTTACAAAGACA 1380
Db 1661 TGCCACAGTAGAATGTCTGTGAACAGAGAGCCCTTGTGGGTCATGTCTTACAAAGACA 1720
Qy 1381 AAATGTGTCTTCTCCTGAACCATGTGTGATAACTTTTACAGAAATGGACTGGATCTATCTG 1440

Db 1721 AAAGTCTGTCTTCTCTGAACCATGTGGATAACTTTTACAGAAATGGACTGGAGTCACTG 1780
Qy 1441 CAAAAGGCTCTTGTAAAGACTGGTTTCTGCAATGACCAACAGCAAGATTTTCTCTC 1500
Db 1781 CAAAAGGCTCTTGTAAAGACTGGTTTCTGCAATGACCAACAGCAAGATTTTCTCTC 1840
Qy 1501 TTGTGATTTCTTTAAAGAAAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Db 1841 TTGTGATTTCTTTAAAGAAAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1900
Qy 1561 ATTCAATTTTATAGCAACAACAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACAACAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1655
Db 1961 GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1995

RESULT 13
BD247178 1997 bp DNA linear PAT 17-JUL-2003
LOCUS BD247178
DEFINITION Use of VEGF-C or VEGF-D gene or protein to prevent restenosis.
ACCESSION BD247178
VERSION BD247178.1 GI:33056948
KEYWORDS JP 2002528420-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Hertzuala,S.Y., Alitalo,K., Hiltunen,M.O., Jeltsch,M.M. and Achen,M.G.
TITLE Use of VEGF-C or VEGF-D gene or protein to prevent restenosis
JOURNAL Patent: JP 2002528420-A 1 03-SEP-2002;
LUDWIG INSTITUTE FOR CANCER RESEARCH, HELSINKI UNIVERSITY LICENSING
LTD OY, SEPPO YLA HERTTUALA
COMMENT OS Homo sapiens (human)
PN JP 2002528420-A/1
PD 03-SEP-2002
PF 26-OCT-1999 JP 2000578021
PR 28-OCT-1998 US 60/105587
PI SEPPO YLA HERTTUALA, KARI ALITALO, MIKKO O HILTUNEN, MARKKU M PI JELTSCH,
PI MARC G ACHEN
PC A61K38/22,A61K31/711,A61K48/00,A61L33/00,A61L33/00,A61M25/00,
PC A61M29/02,A61P41/00,A61P43/00,C07K14/475,C12N15/09,A61K37/24,
PC A61P7/02,A61P41/00,A61L33/00
PC C12N15/00,
PC A61L33/00,A61L33/00
CC Use of VEGF-C or VEGF-D gene or protein to prevent restenosis
FH Key Location/Qualifiers
FT CDS (352)..(1608).
FEATURES
source
1..1997
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GTCTCTCCACCATGCATCGCTGGGCTTCTTCTGTGGGTGTCTCTGTCTGCGGCTG 60
Db 341 GTCTCTCCACCATGCATCGCTGGGCTTCTTCTGTGGGTGTCTCTGTCTGCGGCTG 400
Qy 61 CGTGTCTCCCGGTCTCTCGAGCGCGCGCGCGCGCGCTTCAGTCCGAGATCG 120
Db 401 CGTGTCTCCCGGTCTCTCGAGCGCGCGCGCGCGCGCTTCAGTCCGAGATCG 460
Qy 121 ACCTCTCGAGCGGAGCCCGACGCGGCGGAGCGCGGCTTATGCAAGCAAGATCTGG 180

461	Db	ACCTCTCGGACGGGAGCCGACCGCGGGCAGGCCACGGCTATTGCAAGCAAGAATCTCG	520
181	Qy	AGGAGCAGTTACCGTCTGTCTCCAGTGTAGATGAACTCATGACTGTACTCTACCCAGAAT	240
521	Db	AGGACAGTTACCGTCTGTGTCCAGTGTAGATGAACTCATGACTGTACTCTACCCAGAAT	580
241	Qy	ATTGGAAAAATGTACAAGTGTCACTAAGGAAAGGAGCTGGCAACATAACAGAGAAACAGG	300
581	Db	ATTGGAAAAATGTACAAGTGTCACTAAGGAAAGGAGCTGGCAACATAACAGAGAAACAGG	640
301	Qy	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGACACATTAATAATACAG	360
641	Db	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGACACATTAATAATACAG	700
361	Qy	AGATCTTTGAAAAATATTGATAATGAGTGGGAGAAAGACTCAATCATGCCACGGGAGGTGT	420
701	Db	AGATCTTTGAAAAATATTGATAATGAGTGGGAGAAAGACTCAATCATGCCACGGGAGGTGT	760
421	Qy	GTATAGATGTGGGAGAGGATTTGGAGTCCGACAAACACCTCTTTTAAACCTCCATGTG	480
761	Db	GTATAGATGTGGGAGAGGATTTGGAGTCCGACAAACACCTCTTTTAAACCTCCATGTG	820
481	Qy	TGTCCTGTACAGATGTGGGGTGTCTGCAATAGTAGGGGCTGCAGTGCATGAACACCA	540
821	Db	TGTCCTGTACAGATGTGGGGTGTCTGCAATAGTAGGGGCTGCAGTGCATGAACACCA	880
541	Qy	GCACGAGCTACCTACGACGAGCTTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
881	Db	GCACGAGCTACCTACGACGAGCTTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	940
601	Qy	AACCAGTAAACAATCAGTTTGGCCAAACACACTTCTCTCCGATGCATGCTCAAACTGGATG	660
941	Db	AACCAGTAAACAATCAGTTTGGCCAAACACACTTCTCTCCGATGCATGCTCAAACTGGATG	1000
661	Qy	TTTACAGACAAGTTCAATTCATTATTAGACGTTCCCTGCCAGCAAACTACCAACAGTGT	720
1001	Db	TTTACAGACAAGTTCAATTCATTATTAGACGTTCCCTGCCAGCAAACTACCAACAGTGT	1060
721	Qy	AGGACGGAACAGACCTGCCCCACCAATTACATGTGGAAATATCAATCTGCAGATGCC	780
1061	Db	AGGACGGAACAGACCTGCCCCACCAATTACATGTGGAAATATCAATCTGCAGATGCC	1120
781	Qy	TGGCTCAGGAAGATTTTATGTTTTCCTCGGATGTGGAGATGACTCAACAGATGGATTC	840
1121	Db	TGGCTCAGGAAGATTTTATGTTTTCCTCGGATGTGGAGATGACTCAACAGATGGATTC	1180
841	Qy	ATGACATCTGTGGACCAAAAACAGAGCTGGAATGAAGAGACCTGTGCTGTGTCTGAGAG	900
1181	Db	ATGACATCTGTGGACCAAAAACAGAGCTGGAATGAAGAGACCTGTGCTGTGTCTGAGAG	1240
901	Qy	CGGGGCTTCGGCCGCGCAGCTGTGACCCCAACAAGAACTAGACAGAAAATCTCATGCCAGT	960
1241	Db	CGGGGCTTCGGCCGCGCAGCTGTGACCCCAACAAGAACTAGACAGAAAATCTCATGCCAGT	1300
961	Qy	GTGTCTGTAAAAACAAACTCTTCCCCAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA	1020
1301	Db	GTGTCTGTAAAAACAAACTCTTCCCCAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA	1360
1021	Qy	ACACATGCCAGTGTATGTATTAAGAACCTGCCCCAGAAATCAACCCCTTAATCCTCGAA	1080
1361	Db	ACACATGCCAGTGTATGTATTAAGAACCTGCCCCAGAAATCAACCCCTTAATCCTCGAA	1420
1081	Qy	AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGGAAAGAGTTCC	1140
1421	Db	AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGGAAAGAGTTCC	1480
1141	Qy	ACCACCAACATGACGCTTTACAGACGGCCATGTACGAAACGGCCAGNAGCTTGTGAGC	1200
1481	Db	ACCACCAACATGACGCTTTACAGACGGCCATGTACGAAACGGCCAGNAGCTTGTGAGC	1540
1201	Qy	CAGGATTTTCAATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATTGGCAAAAGACAC	1260

Db	1541	CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTGCCCTTCATATTGGAAGAACCAC	160			
Qy	1261	AAATGAGCTAAGATTTGTACTGTTTTCCAGTTCATCGAATTTTCTATTATGAAAACTGTGT	1320			
Db	1601	AAATGAGCTAAGATTTGTACTGTTTTCCAGTTCATCGAATTTTCTATTATGAAAACTGTGT	1660			
Qy	1321	TGCCACAGTGAAGACTGTCGTGAACAGAGAGACCCCTGTGGGTCCATGCTTAAACAAGACA	1380			
Db	1661	TGCCACAGTGAAGACTGTCGTGAACAGAGAGACCCCTGTGGGTCCATGCTTAAACAAGACA	1720			
Qy	1381	AAAGTCTGCTTTCCCTGAACCATGTGGATAACTTTTACAGAAATGGACTGGAGCTCATCTG	1440			
Db	1721	AAAGTCTGCTTTCCCTGAACCATGTGGATAACTTTTACAGAAATGGACTGGAGCTCATCTG	1780			
Qy	1441	CAAAAGCCCTCTGTTAAAGACTGGTTTTCTGCCAATGACCAACAGACCAAGATTTTCCTC	1500			
Db	1781	CAAAAGCCCTCTGTTAAAGACTGGTTTTCTGCCAATGACCAACAGACCAAGATTTTCCTC	1840			
Qy	1501	TTGTGATTTCTTTAAAGAAATGACTATATATAATTTATTTCCACTAAAAATATTGTTTCTGC	1560			
Db	1841	TTGTGATTTCTTTAAAGAAATGACTATATATAATTTATTTCCACTAAAAATATTGTTTCTGC	1900			
Qy	1561	ATTCAATTTTATAGCAACAAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT	1620			
Db	1901	ATTCAATTTTATAGCAACAAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT	1960			
Qy	1621	GCAAAATATCTTTAAATAATAAAATGAAATTTGTATT	1655			
Db	1961	GCAAAATATCTTTAAATAATAAAATGAAATTTGTATT	1995			
RESULT 14						
AR201986						
LOCUS	AR201986	1997 bp	DNA linear PAT 20-APR-2002			
DEFINITION	Sequence 7 from patent US 6361946.					
ACCESSION	AR201986					
VERSION	AR201986.1 GI:20256525					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1997)					
AUTHORS	Altaio, K. and Joukov, V.					
TITLE	Vascular endothelial growth factor C (VEGF-C) protein and gene, mutants thereof, and uses thereof					
JOURNAL	Patent: US 6361946-A 7 26-MAR-2002;					
FEATURES	Location/Qualifiers					
	1..1997					
source	/organism="unknown"					
ORIGIN	/mol_type="unassigned DNA"					
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;						
Best Local Similarity 99.9%; Pred. No. 0;						
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;						
Qy	1	GTCCTTCCACATGACATCGCTGGCTTCTTCTCTGTGGCGTGTCTCTGTCTGCTCGCCGCTG	60			
Db	341	GTCCTTCCACATGACATCGCTGGCTTCTTCTCTGTGGCGTGTCTCTGTCTGCTCGCCGCTG	400			
Qy	61	CGCTGCTCCCGGGTCTTCGCGAGCGCCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG	120			
Db	401	CGCTGCTCCCGGGTCTTCGCGAGCGCCGCGCGCGCGCGCTTCGAGTCCGAGCTCG	460			
Qy	121	ACCTCTCGACCGGAGCCCGACCGCGGCGAGGCCACCGCTTATGCAAGCAAGATCTCG	180			
Db	461	ACCTCTCGACCGGAGCCCGACCGCGGCGAGGCCACCGCTTATGCAAGCAAGATCTCG	520			
Qy	181	AGGACGAGTACGGTCTGTGTCAGTGATGTAATCATGACTGTACTCTTACCCAGAAAT	240			
Db	521	AGGACGAGTACGGTCTGTGTCAGTGATGTAATCATGACTGTACTCTTACCCAGAAAT	580			
Qy	241	ATTGGAAATGTACAAGTGTGCTTACGCTTAAGGAAAGGCGGTGGCAACATACGAGAACGAG	300			

Db	701	AGATCTTTGAAAGTATTGATTAATGAGTGGAGAAAGACTCAATGTCATGCCGAGGAGTGT	760
Qy	421	GTATAGATGTGGGAAGAGTTTGGAGTCGGACGCAACACACCTTCTTTAAACCTCCATGTG	480
Db	761	GTATAGATGTGGGAAGAGTTTGGAGTCGGACGCAACACACCTTCTTTAAACCTCCATGTG	820
Qy	481	TGTCGGTCTACAGATGTGGGGTTTCTGTGCAATAGTGGGGCTGCAGTGCATGAACACCA	540
Db	821	TGTCGGTCTACAGATGTGGGGTTTCTGTGCAATAGTGGGGCTGCAGTGCATGAACACCA	880
Qy	541	GCACGAGTACTCAGCAAGACGTTATTTGAAATTACAGTGCCTCTCTCTCAAGGCCCA	600
Db	881	GCACGAGTACTCAGCAAGACGTTATTTGAAATTACAGTGCCTCTCTCTCAAGGCCCA	940
Qy	601	AACCAAGTAAACATCAGTTTTCGCAATCACACTTCTCTGCCATGCATGCTCTAACTGGATG	660
Db	941	AACCAAGTAAACATCAGTTTTCGCAATCACACTTCTCTGCCATGCATGCTCTAACTGGATG	1000
Qy	661	TTTACAGACAAGTTCAITTCATTTATAGACGTTCCCTGCCAGCAACACTACCAGTGTG	720
Db	1001	TTTACAGACAAGTTCAITTCATTTATAGACGTTCCCTGCCAGCAACACTACCAGTGTG	1060
Qy	721	AGGACGCAACAAGACCTGCCCCCAATACATGTGGAATAATCACATCTGCAGATGCC	780
Db	1061	AGGACGCAACAAGACCTGCCCCCAATACATGTGGAATAATCACATCTGCAGATGCC	1120
Qy	781	TGGCTCAGAGAATTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC	840
Db	1121	TGGCTCAGAGAATTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC	1180
Qy	841	ATGACATCTGTGGACCAACAGAGCTGGATGAAGACCTGTGAGTGTCTGTCAGAG	900
Db	1181	ATGACATCTGTGGACCAACAGAGCTGGATGAAGACCTGTGAGTGTCTGTCAGAG	1240
Qy	901	CGGGCTTTCGGCTGCCAGCTGTGGAACCCCAAGAACTAGACAGAACTCATGCCAGT	960
Db	1241	CGGGCTTTCGGCTGCCAGCTGTGGAACCCCAAGAACTAGACAGAACTCATGCCAGT	1300
Qy	961	GTGCTGTGTAACAACTCTTCCAGCAGCAATGTGGGGCCCAACGGAGATTTGATGAA	1020
Db	1301	GTGCTGTGTAACAACTCTTCCAGCAGCAATGTGGGGCCCAACGGAGATTTGATGAA	1360
Qy	1021	ACACATGCCAGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTCAATCCTGGAA	1080
Db	1361	ACACATGCCAGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTCAATCCTGGAA	1420
Qy	1081	AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAGAGTTCC	1140
Db	1421	AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAGAGTTCC	1480
Qy	1141	ACCACCAACATGCAGCTGTACAGACGGCCATGTACGAACCCGCGAGAGCTTGTGAGC	1200
Db	1481	ACCACCAACATGCAGCTGTACAGACGGCCATGTACGAACCCGCGAGAGCTTGTGAGC	1540
Qy	1201	CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTCCTTCATATTGGCAAGACAC	1260
Db	1541	CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTCCTTCATATTGGCAAGACAC	1600
Qy	1261	AAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTTCTATATGGAATACTGT	1320
Db	1601	AAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTTCTATATGGAATACTGT	1660
Qy	1321	TGCCACAGTAGAATCTGTGTGAACAGAGAGACCTTGTGGTCCATGCTAACCAAGACA	1380
Db	1661	TGCCACAGTAGAATCTGTGTGAACAGAGAGACCTTGTGGTCCATGCTAACCAAGACA	1720
Qy	1381	AAAGTCTGCTTCTGACCACTGTGATTAATTTACAGAAATGGACTGGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTCTGACCACTGTGATTAATTTACAGAAATGGACTGGAGCTCATCTG	1780
Qy	1441	CAAAAGGCCCTGTGTAAGACTGGTTTCTGCCAATGACCAACAGCCAGATTTTCTC	1500

Search completed: February 25, 2004, 20:57:49
Job time : 6423 secs

Db	1781	CAAAAGGCCCTGTGTAAGACTGGTTTCTGCCAATGACCAACAGCCAGAGATTTTCTC	1840
Qy	1501	TTGTGATTTCTTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC	1560
Db	1841	TTGTGATTTCTTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC	1900
Qy	1561	ATTCAATTTTATAGCAACAAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT	1620
Db	1901	ATTCAATTTTATAGCAACAAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT	1960
Qy	1621	GCAAAATATGTTTAAATATAAAATGAAATTTGTATT	1655
Db	1961	GCAAAATATGTTTAAATATAAAATGAAATTTGTATT	1995

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 16:56:30 ; Search time 653 Seconds
(without alignments)
10890.480 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gctctccaccatgcactcg.....ttataaaaaaaaaaaaaaa 1674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	100.0	1674	2	AAX84837 Full leng
2	1674	100.0	1674	2	AAX10523 CDNA enco
3	1674	100.0	1674	3	AA52080 Vascular
4	1674	100.0	1674	4	AA91004 Human VEG
5	1674	100.0	1674	6	ABX10034 Human vas
6	1674	100.0	1674	7	ABT17108 Human VEG
7	1674	100.0	1674	7	ABQ76966 Human VEG
8	1674	100.0	1674	7	AAD49522 Human vas
9	1667.6	99.6	1674	2	AAT51371 Human vas
10	1664.6	99.4	2031	2	AAT59929 Human vas
11	1651.8	98.7	1997	2	AAT84276 Human vas
12	1651.8	98.7	1997	2	AAV52576 Human vas
13	1651.8	98.7	1997	3	AA000339 Human vas
14	1651.8	98.7	1997	3	AAC62406 Human VEG
15	1651.8	98.7	1997	4	AAC68953 Human VEG
16	1651.8	98.7	1997	4	AA512820 DNA enco
17	1651.8	98.7	1997	9	ADD08947 Human VEG
18	1651.8	98.7	2015	9	ADC64993 Human VEG
19	1651.8	98.7	2879	3	AA000353 CMV-VEGF-
20	1607.8	96.0	2321	2	AAT68811 Human fce
21	1526	91.2	1526	2	AAX84838 Truncated
22	1526	91.2	1526	2	AA510524 CDNA enco
23	1526	91.2	1526	4	AA91010 Human VEG

24	1526	91.2	1526	6	ABK10035 Human vas
25	1520.2	90.8	1525	3	AA52081 Truncated
26	1520.2	90.8	1525	7	ABT17109 Human VEG
27	1520.2	90.8	1525	7	ABQ76967 Human tru
28	1520.2	90.8	1525	7	AAD49523 Human vas
29	1517	90.6	1525	7	AAT03950 DNA enco
30	1266.6	75.7	5283	6	ABK10062 Expressio
31	1256.8	75.1	1260	6	ABV78239 Human VEG
32	1256.8	75.1	1260	6	ABZ35815 Human VEG
33	1256.8	75.1	1260	6	ABX10058 Human VEG
34	1256.8	75.1	1260	6	ABY91780 Human pol
35	1182.2	70.6	1836	2	AAT84277 Mouse Flt
36	1182.2	70.6	1836	2	AAV52577 Mouse vas
37	1076.4	64.3	1596	7	ABT42320 Toxicity
38	1076.4	64.3	1596	9	AD53485 Primary r
39	804.2	48.0	1741	2	AAT84300 Quail Flt
40	804.2	48.0	1741	2	AAV52578 Quail vas
41	298	17.8	239	4	AAT59930 EST HSC1W
42	278.2	16.6	397	4	AA512843 DNA seque
43	276.2	16.5	1215	7	AB856721 VEGF-C co
44	231.8	13.8	331	4	AA512888 DNA seque
45	227	13.6	331	4	AA512889 DNA seque

ALIGNMENTS

RESULT 1

AAX84837
ID AAX84837 standard; cDNA; 1674 BP.

XX AAX84837;

DT 22-SEP-1999 (first entry)

XX Full length human VEGF2 coding sequence.

DE Full length human VEGF2 coding sequence.

KW VEGF2; vascular endothelial growth factor 2; angiogenesis; bone damage;

KW endothelial cell proliferation; tissue damage; therapy; ds.

XX Homo sapiens.

XX US5932540-A.

XX 03-AUG-1999.

XX 24-DEC-1997; 97US-00999811.

XX 08-MAR-1994; 94US-00207550.

XX 06-JUN-1995; 95US-00465968.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Cao L, Rosen CA, Hu J;

XX WPI; 1999-443606/37.

XX P-PSDB; AAY22320.

XX Vascular endothelial growth factor 2 for wound healing and vascular repair.

XX Disclosure; Fig 1; 49pp; English.

XX This sequence encodes the vascular endothelial growth factor 2 (VEGF2),

XX of the invention. The isolated polypeptide is useful for stimulating

XX angiogenesis, by promoting the proliferation of endothelial cells, for

XX the treatment of a wound, or for the treatment of tissue or bone damage

SQ Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 1674; DB 2; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTCCTTCCACCATGCACTCGCTGGGCTTCTTCTCTGTGCGGTGTTCTCTGTCTCGCGCTG 60
Db 1 GTCCTTCCACCATGCACTCGCTGGGCTTCTTCTCTGTGCGGTGTTCTCTGTCTCGCGCTG 60
QY 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCGCGCGCGCGCGCGCCCTTCAGTCCGAGCTCG 120
Db 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCGCGCGCGCGCGCGCCCTTCAGTCCGAGCTCG 120
QY 121 ACCTCTCGAGCGGAGCGCGAGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGAGCGGAGCGCGAGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
QY 181 AGAGCAGTGTACGGTCTGTGTCAGTGTAGATGAATCACTGATGACTGTACTCTACCCAGAAT 240
Db 181 AGAGCAGTGTACGGTCTGTGTCAGTGTAGATGAATCACTGATGACTGTACTCTACCCAGAAT 240
QY 241 ATTGGAAAATGTACAAGTGTACCTAAGGAAGAGGCTGGCAACATTAACAGAGAACAGG 300
Db 241 ATTGGAAAATGTACAAGTGTACCTAAGGAAGAGGCTGGCAACATTAACAGAGAACAGG 300
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGACACATTAATATACAG 360
Db 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGACACATTAATATACAG 360
QY 361 AGATCTTGAAGATATTGATTAATGAGTGGAGAGAGCTCAATGCTCATGCCACGGAGGTGT 420
Db 361 AGATCTTGAAGATATTGATTAATGAGTGGAGAGAGCTCAATGCTCATGCCACGGAGGTGT 420
QY 421 GTATAGATGTGGGAGGAGTGTGGAGTGCAGCAAAACACCTCTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGAGGAGTGTGGAGTGCAGCAAAACACCTCTTTAAACCTCCATGTG 480
QY 481 TGTCCGCTACAGATGTGGGCTTGTGCAATAGTGGAGGCTGCAGTGCATGCAACACCA 540
Db 481 TGTCCGCTACAGATGTGGGCTTGTGCAATAGTGGAGGCTGCAGTGCATGCAACACCA 540
QY 541 GCACGAGTACCTCAGCAAGACGTATTGTAATTTACAGTGCCTCTCTCAAGGCCCA 600
Db 541 GCACGAGTACCTCAGCAAGACGTATTGTAATTTACAGTGCCTCTCTCAAGGCCCA 600
QY 601 AACGAGTAAACATGAGTTTGCATTCACATTCACACTTCTCCGCGATGCATCTTAACCTGGATG 660
Db 601 AACGAGTAAACATGAGTTTGCATTCACATTCACACTTCTCCGCGATGCATCTTAACCTGGATG 660
QY 661 TTTACAGACAAGTTCATTCATTTAGACGTTTCCCTGCCAGCAACACTACCACAGTGC 720
Db 661 TTTACAGACAAGTTCATTCATTTAGACGTTTCCCTGCCAGCAACACTACCACAGTGC 720
QY 721 AGCAGCGAAACAAGACTGCCCAACCAATTTACATGTGGAAATATCACTGTGAGATGCC 780
Db 721 AGCAGCGAAACAAGACTGCCCAACCAATTTACATGTGGAAATATCACTGTGAGATGCC 780
QY 781 TGGCTCAGGAGATTTTATGTTTCTTCGATGCTGGAGATGACTCAACAGATGGATTC 840
Db 781 TGGCTCAGGAGATTTTATGTTTCTTCGATGCTGGAGATGACTCAACAGATGGATTC 840
QY 841 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCAAGTGTGTCTGAGAG 900
Db 841 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCAAGTGTGTCTGAGAG 900
QY 901 CGGGCTTCGGCTCGCCAGTGTGGCCCGCACAGAACTAGACAGAACTCATGCGCAGT 960
Db 901 CGGGCTTCGGCTCGCCAGTGTGGCCCGCACAGAACTAGACAGAACTCATGCGCAGT 960
QY 961 GTGCTGTAAAAAACAACCTCTTCCAGCCAAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
Db 961 GTGCTGTAAAAAACAACCTCTTCCAGCCAAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
QY 1021 ACATATGCCATGTGTATGTAAAAAGAACCTGCCCGCAAGATCAACCCCTAAATCCTGGAA 1080
Db 1021 ACATATGCCATGTGTATGTAAAAAGAACCTGCCCGCAAGATCAACCCCTAAATCCTGGAA 1080
```

```
QY 1081 AATGTGCTCTGTAATGTACAGAAAGTCCACAGAAATGCTTGTATAAGGAAGAGTTCC 1140
Db 1081 AATGTGCTCTGTAATGTACAGAAAGTCCACAGAAATGCTTGTATAAGGAAGAGTTCC 1140
QY 1141 ACCACCAAAACATGAGCTGTTACAGACGGCGCATGTAGAACCCGCCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAAAACATGAGCTGTTACAGACGGCGCATGTAGAACCCGCCAGAGGCTTGTGAGC 1200
QY 1201 CAGATTTTATATAGTGAAGAGTGTGCTGTTGTGCTCCCTTCATATTTGCAAGACCAC 1260
Db 1201 CAGATTTTATATAGTGAAGAGTGTGCTGTTGTGCTCCCTTCATATTTGCAAGACCAC 1260
QY 1261 AAATGAGCTAAGATTTGCTGCTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1320
Db 1261 AAATGAGCTAAGATTTGCTGCTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1320
QY 1321 TGCCACAGTGAAGTGTCTGTGAACAGAGAGACCTTTGTGGTCCATGCTAAACAAGACA 1380
Db 1321 TGCCACAGTGAAGTGTCTGTGAACAGAGAGACCTTTGTGGTCCATGCTAAACAAGACA 1380
QY 1381 AAAGTCTGCTTCTTCCAGCAACATGTGGATAAATTTACAGAAATGGAGCTCATCTG 1440
Db 1381 AAAGTCTGCTTCTTCCAGCAACATGTGGATAAATTTACAGAAATGGAGCTCATCTG 1440
QY 1441 CAAAAGCCCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCCCTC 1500
Db 1441 CAAAAGCCCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCCCTC 1500
QY 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTTCCACTAAATAATTTGTTCTGC 1560
Db 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTTCCACTAAATAATTTGTTCTGC 1560
QY 1561 ATTCAATTTTATAGCAACAAATTTGTAAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAAATTTGTAAACTCACTGTGATCAATATTTTATATCAT 1620
QY 1621 GCAAAATATGTTTAAATTAATAAATTTGTAATTTTATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATTAATAAATTTGTAATTTTATAAAAAA 1674

RESULT 2
AAZ10523
ID AAZ10523 standard; cDNA; 1674 BP.
XX
AC AAZ10523;
XX
DT 16-NOV-1999 (first entry)
XX
DE cDNA encoding vascular endothelial growth factor-2 (VEGF-2).
XX
KW Human vascular endothelial growth factor-2; VEGF-2;
KW vascular endothelial cell growth; endothelial cell migration;
KW angiogenesis; blood pressure; blood flow; immune system disorder;
KW immune cell; cancer; autoimmune disorder; blood protein disorder;
KW ataxia telangiectasia; common variable immunodeficiency;
KW Digorge syndrome; HIV infection; HTLV-BLV infection;
KW leukocyte adhesion deficiency syndrome; lymphopenia;
KW phagocyte bactericidal dysfunction; severe combined immunodeficiency;
KW Wiskott-Aldrich disorder; anemia; thrombocytopenia; hemoglobinuria;
KW allergy; asthma; allergic asthma; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 12..1271
FT /*tag= a
FT /product= "VEGF-2"
FT sig_peptide 12..80
FT /*tag= b
FT mat_peptide 81..1268
FT /*tag= c
XX
```


Db 1441 CAAAAGGCGCTCTGTAAAGACTGGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1500
 Qy 1501 TTGTGATTTCTTTAAAGAAAGACTATATAATTTTCCACTAAATAATTTTCTGC 1560
 Db 1501 TTGTGATTTCTTTAAAGAAAGACTATATAATTTTCCACTAAATAATTTTCTGC 1560
 Qy 1561 ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
 Db 1561 ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
 Qy 1621 GCAAAATATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1674
 Db 1621 GCAAAATATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1674

RESULT 3
 AA52080
 ID AAA52080 standard; cDNA; 1674 BP.
 XX AA52080;
 AC XX
 DT 22-DEC-2000 (first entry)
 XX Vascular endothelial growth factor-2 (VEGF-2) coding sequence.
 DE Vascular endothelial growth factor 2; VEGF-2; retina; angiogenesis;
 XX Vascular endothelial growth factor 2; VEGF-2; retina; angiogenesis;
 KW treatment; injury; degeneration; photoreceptors; eye; angiod streaks;
 KW retinitis; pigmentosa; age-related macular degeneration;
 KW diabetic retinopathy; human; ds.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 CDS 12..1271
 FT /*tag= a
 FT /product= "Vascular endothelial growth factor-2"
 XX WO200045835-A1.
 XX
 PD 10-AUG-2000.
 XX
 PF 07-FEB-2000; 2000WO-US003047.
 XX
 PR 08-FEB-1999; 99US-0119179P.
 PR 12-FEB-1999; 99US-0119926P.
 PR 03-JUN-1999; 99US-0127726P.
 PR 22-DEC-1999; 99US-0171505P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Alderson R, Melder R, Roschke V, Ruben SM;
 PI
 XX WPI: 2000-532862/48.
 DR P-PSDB; AA197144.
 DR
 XX Treating injury or degeneration of photoreceptors comprises administering
 PT to a subject vascular endothelial growth factor 2 (VEGF-2).
 XX
 PS Claim 24; Fig 1a-e; 252pp; English.
 XX
 CC Administration of vascular endothelial growth factor 2 (VEGF-2) to a
 CC patient can be used for treating injury or degeneration of photoreceptors
 CC associated with e.g. angiod streaks, retinitis pigmentosa, age-related
 CC macular degeneration, diabetic retinopathy, etc. VEGF-2 promotes
 CC angiogenesis, the formation of new blood vessels in the retina
 XX
 XX Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 1674; DB 3; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCACTGGTGGGCTTCTTCTCTGGGTGTTCTCTGCTCGCGCTG 60
 Db 1 GTCTTCCACCATGCACTGGTGGGCTTCTTCTCTGGGTGTTCTCTGCTCGCGCTG 60
 Qy 61 CGTGTCTCCCGGCTCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
 Db 61 CGTGTCTCCCGGCTCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
 Qy 121 ACCTCTCGGAGCTG 180
 Db 121 ACCTCTCGGAGCTG 180
 Qy 181 AGGAGCAGTTTACGCTCTGCTCCAGTGTAGATGAATCTCATGCTGTACTTACCCAGAAT 240
 Db 181 AGGAGCAGTTTACGCTCTGCTCCAGTGTAGATGAATCTCATGCTGTACTTACCCAGAAT 240
 Qy 241 ATTGGAATAATGTACAAGTGTCACTAAGAAAGAGGCTGGCAACATACAGAGACAGG 300
 Db 241 ATTGGAATAATGTACAAGTGTCACTAAGAAAGAGGCTGGCAACATACAGAGACAGG 300
 Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCGAGCACATTATAATACAG 360
 Db 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCGAGCACATTATAATACAG 360
 Qy 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
 Db 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
 Qy 421 GTATAGATGTGGGAGAGAGTTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 480
 Db 421 GTATAGATGTGGGAGAGAGTTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 480
 Qy 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGGGGGCTGCAATGATGAACACCA 540
 Db 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGGGGGCTGCAATGATGAACACCA 540
 Qy 541 GCACGAGTACCTCAGCAAGACGTTTATTTGAAATACAGTGCCTCTCTCTCAAGGCCCA 600
 Db 541 GCACGAGTACCTCAGCAAGACGTTTATTTGAAATACAGTGCCTCTCTCTCAAGGCCCA 600
 Qy 601 AACCAGTAACCAATCAGTGTTCCTGCAATCAGTGTTCCTGCAATGATGATGATGATG 660
 Db 601 AACCAGTAACCAATCAGTGTTCCTGCAATCAGTGTTCCTGCAATGATGATGATGATG 660
 Qy 661 TTTACAGACAAAGTTCAATTTCAATTTAGAGTTCCTGCGGAGCAACACTACACAGTGC 720
 Db 661 TTTACAGACAAAGTTCAATTTCAATTTAGAGTTCCTGCGGAGCAACACTACACAGTGC 720
 Qy 721 AGGAGCGCAACAGACCTGCCCACTACATGATGATGATGATGATGATGATGATGATG 780
 Db 721 AGGAGCGCAACAGACCTGCCCACTACATGATGATGATGATGATGATGATGATGATG 780
 Qy 781 TGGCTCAGGAAAGATTTTATGTTTCTCGGATGCTGGAGTGTGATGATGATGATGATG 840
 Db 781 TGGCTCAGGAAAGATTTTATGTTTCTCGGATGCTGGAGTGTGATGATGATGATGATG 840
 Qy 841 ATGACATCTGGGAGCAACAGAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 900
 Db 841 ATGACATCTGGGAGCAACAGAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 900
 Qy 901 CGGGCTTTCGGCTGCCAGCTGTGAGCCCAACAAAGAACTAGACAGAACTATGCGCAGT 960
 Db 901 CGGGCTTTCGGCTGCCAGCTGTGAGCCCAACAAAGAACTAGACAGAACTATGCGCAGT 960
 Qy 961 GTGTCTGTAAAAAAMAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGATTTTATGAAA 1020
 Db 961 GTGTCTGTAAAAAAMAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGATTTTATGAAA 1020
 Qy 1021 ACACATGCGCAGTGTGATGTAATAAGAACTGCCCCAGAAAATCAACCCCTTAATCTGAAA 1080
 Db 1021 ACACATGCGCAGTGTGATGTAATAAGAACTGCCCCAGAAAATCAACCCCTTAATCTGAAA 1080
 Qy 1081 AATGTCCTGTGAAATGTTACAGAAAAGTCCAGAAAATGCTTGTAAAGAGAAAGATTC 1140

Db 1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTAAAGGAAAGAGTCTCC 1140
QY 1141 ACCACCAAAACATGACAGCTGTTACAGACGGCCATGTACGAACCCGCCAGAGGCTTTGTGAGC 1200
Db 1141 ACCACCAAAACATGACAGCTGTTACAGACGGCCATGTACGAACCCGCCAGAGGCTTTGTGAGC 1200
QY 1201 CAGGATTTTCATATAGTGAAGAAGTGTGCTGTTGTCCTTCATATTTGCGCAAGACAC 1260
Db 1201 CAGGATTTTCATATAGTGAAGAAGTGTGCTGTTGTCCTTCATATTTGCGCAAGACAC 1260
QY 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTCTATATATGAAAACTGTGT 1320
Db 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTCTATATATGAAAACTGTGT 1320
QY 1321 TGCCACAGTAGAATCTGCTGTGAACAGAGAGACCCCTTGGGTCCATGTCTAACAAAGACA 1380
Db 1321 TGCCACAGTAGAATCTGCTGTGAACAGAGAGACCCCTTGGGTCCATGTCTAACAAAGACA 1380
QY 1381 AAGATCTGCTTTTCCCTGAACCATGTGGATACTTTACAGAAATGGAGCTGGAGCTCATCTG 1440
Db 1381 AAGATCTGCTTTTCCCTGAACCATGTGGATACTTTACAGAAATGGAGCTGGAGCTCATCTG 1440
QY 1441 CAAAAGCCCTCTTTGTAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTTC 1500
Db 1441 CAAAAGCCCTCTTTGTAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTTC 1500
QY 1501 TTGTGATTTCTTTAAAGAGATGACTATATATTTTCCACTAAATAATTTGTTCTGC 1560
Db 1501 TTGTGATTTCTTTAAAGAGATGACTATATATTTTCCACTAAATAATTTGTTCTGC 1560
QY 1561 ATTCAATTTTATAGCAACAATTTGGTAAATCTCACTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAATTTGGTAAATCTCACTGTGATCAATATTTTATATCAT 1620
QY 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTTATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTTATAAAAAA 1674

RESULT 4

ID AAA91004 standard; DNA; 1674 BP.
AC
XX
AC
XX
AC
XX
05-APR-2001 (first entry)
XX
XX
Human VEGF-B coding sequence.
XX
XX
Human; angiogenic protein; wound healing; vascular tissue repair;
KW peripheral arterial disease; critical limb ischaemia; coronary disease;
KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KW infectious disease; neurodegeneration;
KW vascular endothelial growth factor-B; VEGF-B; ds.
XX
OS Homo sapiens.
XX
XX
Key Location/Qualifiers
FH 12..1271
CDS /*tag= a
FT /product= "VEGF-B"
XX
XX
WO200075163-A1.
XX
XX 14-DEC-2000.
XX
XX 01-JUN-2000; 2000HO-US014925.
XX
XX 03-JUN-1999; 99US-0137796P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Hu J, Cao L;
XX WPI; 2001-071057/08.
DR P-PSDB; RAY97570.
XX
PT New nucleic acid encoding angiogenic proteins, useful e.g. for promoting
PT healing of wounds and treating peripheral arterial disease, critical limb
PT ischemia or coronary disease.
XX
PS Claim 1; Fig 1; 24app; English.
XX
CC This sequence encodes vascular endothelial growth factor-B (VEGF-B),
CC which is an angiogenic protein of the invention. The angiogenic proteins
CC and the DNA sequences encoding them, are used to prevent, treat or
CC ameliorate disease and to detect diseases, or susceptibility, by
CC detecting mutations or the presence or amount of angiogenic protein
CC expression. Particularly they are used to stimulate wound healing, growth
CC of damaged bone and tissue, and for repair of vascular tissue, especially
CC peripheral arterial disease, critical limb ischaemia or coronary disease.
CC Antagonists of the sequences are used to inhibit angiogenesis in tumours
CC and to treat inflammation (where associated with increased vascular
CC permeability), diabetic retinopathy, rheumatoid arthritis or psoriasis.
CC Agonists are also useful for stimulating (lymph)angiogenesis. The
CC proteins are also used to identify specific binding agents (potential
CC therapeutic agents) and to raise antibodies. The antibodies are useful as
CC therapeutic (ant)agonists; for detection, purification and targeting of
CC proteins for in vivo or in vitro diagnosis (including imaging) or for
CC therapy (including when linked to e.g. a label or cytotoxin); and for
CC immunotyping of cells, e.g. for detecting minimal residual disease or
CC haematopoietic progenitor/stem cells. It is also contemplated that the
CC sequences might be useful for treating a very wide range of other
CC disorders, e.g. autoimmune diseases; allergy; cancer; infectious diseases
CC (viral, bacterial, fungal or parasitic); neurodegeneration, also as
CC chemotactic agents or for stimulating regeneration of the nervous system
CC etc
XX
XX Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1674; DB 4; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCCTTCCACCATGACATCGCTGGGCTTCTTCTGTGGCGTGTCTGTCTGCGCGCTG 60
Db 1 GTCCTTCCACCATGACATCGCTGGGCTTCTTCTGTGGCGTGTCTGTCTGCGCGCTG 60
QY 61 CGCTGCTCCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCGCTTCGAGTCGGACTCG 120
Db 61 CGCTGCTCCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCGCTTCGAGTCGGACTCG 120
QY 121 ACCTTCGGACCGGAGCCCGACGCGGCGGCGCGCGCTTATCAAGCAAGATCTGG 180
Db 121 ACCTTCGGACCGGAGCCCGACGCGGCGGCGCGCGCTTATCAAGCAAGATCTGG 180
QY 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTAGTAAGTCTATCTTACCAGAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTAGTAAGTCTATCTTACCAGAT 240
QY 241 ATTGGAATGTACAAGTGTCTAGTAAAGAGGCGTGGCAACATACAGAACAGCAGG 300
Db 241 ATTGGAATGTACAAGTGTCTAGTAAAGAGGCGTGGCAACATACAGAACAGCAGG 300
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCAGCACATTATAACAG 360
Db 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCAGCACATTATAACAG 360
QY 361 AGATCTTGAAGAGTATTGATATATGAGTGAAGAAAGTCAATGATGCGCGGAGGTGT 420
Db 361 AGATCTTGAAGAGTATTGATATATGAGTGAAGAAAGTCAATGATGCGCGGAGGTGT 420
QY 421 GTATAGATCTGGGAGGAGTTTGGAGTCGCGACAAAACACCTTCTTTAAACCTCCATGTG 480

Db	421	GTATAGTGTGGGAGGAGTTTGAGTGGGACAAACACACCTCTCTTAAACCTCCATGTG	480
Qy	481	TGTCGGTCTACAGATGTGGGGTGTCTGCATATAGTAGGGGCTGCAGTGCATGAACACCA	540
Db	481	TGTCGGTCTACAGATGTGGGGTGTCTGCATATAGTAGGGGCTGCAGTGCATGAACACCA	540
Qy	541	GCACGAGTACCTCAGCAAGAGCTTATTTGAAATACAGTGCCTCTCTCAAGGCCCA	600
Db	541	GCACGAGTACCTCAGCAAGAGCTTATTTGAAATACAGTGCCTCTCTCAAGGCCCA	600
Qy	601	AACCAAGTAACTAGTTTTCCTCCATACACTCTCTCCGATGCTAAATCTGGATG	660
Db	601	AACCAAGTAACTAGTTTTCCTCCATACACTCTCTCCGATGCTAAATCTGGATG	660
Qy	661	TTTACAGACAGTTCATTCCTATTTAGAGCTTCCCTGCAGCAACACTACCAAGTGC	720
Db	661	TTTACAGACAGTTCATTCCTATTTAGAGCTTCCCTGCAGCAACACTACCAAGTGC	720
Qy	721	AGGACGCAACAGAGCTGCCCAATACATGTGGAATAATACATCTGCAGATGCC	780
Db	721	AGGACGCAACAGAGCTGCCCAATACATGTGGAATAATACATCTGCAGATGCC	780
Qy	781	TGCTCAGCAAGATTTATGTTTCTCGATGCTGGAGTCACTCAACAGATGATCC	840
Db	781	TGCTCAGCAAGATTTATGTTTCTCGATGCTGGAGTCACTCAACAGATGATCC	840
Qy	841	ATGACATCTGTGACCAACAGAGCTGGATGAAGAGACCTGTGAGTGTGTCTCAGAG	900
Db	841	ATGACATCTGTGACCAACAGAGCTGGATGAAGAGACCTGTGAGTGTGTCTCAGAG	900
Qy	901	CGGGGCTTGGCTCGCAGCTGTGAGCCCAACAGAACTAGACAGAACTATCCAGT	960
Db	901	CGGGGCTTGGCTCGCAGCTGTGAGCCCAACAGAACTAGACAGAACTATCCAGT	960
Qy	961	GTGCTGTAAAACAACTCTTCCCGAGCAATGTGGGGCCCAACGAGATTTGATGAA	1020
Db	961	GTGCTGTAAAACAACTCTTCCCGAGCAATGTGGGGCCCAACGAGATTTGATGAA	1020
Qy	1021	ACACATGCCAGTGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTAAATCTCGAA	1080
Db	1021	ACACATGCCAGTGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTAAATCTCGAA	1080
Qy	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC	1140
Db	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC	1140
Qy	1141	ACCACCAACATGCAGCTGTGTACAGACGGCCATGTACGAACCGCCAGAGAGCTGTGAGC	1200
Db	1141	ACCACCAACATGCAGCTGTGTACAGACGGCCATGTACGAACCGCCAGAGAGCTGTGAGC	1200
Qy	1201	CAGATTTTCATATAGTGAAGAGTGTGCTGCTCCATGATTTGATGAAATCTGTG	1260
Db	1201	CAGATTTTCATATAGTGAAGAGTGTGCTGCTCCATGATTTGATGAAATCTGTG	1260
Qy	1261	AAATGAGTAAAGTGTGCTGCTCCATGATTTGATGAAATCTGTG	1320
Db	1261	AAATGAGTAAAGTGTGCTGCTCCATGATTTGATGAAATCTGTG	1320
Qy	1321	TGCCACAGTAAAGTGTGCTGCTCCATGATTTGATGAAATCTGTG	1380
Db	1321	TGCCACAGTAAAGTGTGCTGCTCCATGATTTGATGAAATCTGTG	1380
Qy	1381	AAAGTGTCTTCTGCAACCTGCTGATTAACCTTACAGAAATGGAGTGCATCTG	1440
Db	1381	AAAGTGTCTTCTGCAACCTGCTGATTAACCTTACAGAAATGGAGTGCATCTG	1440
Qy	1441	CAAAAGGCTCTTCTGCAACCTGCTGATTAACCTTACAGAAATGGAGTGCATCTG	1500
Db	1441	CAAAAGGCTCTTCTGCAACCTGCTGATTAACCTTACAGAAATGGAGTGCATCTG	1500
Qy	1501	TTGTGATTTCTTAAAGATGACTATATATTTTCCACTAAAATATTTGTTCTGC	1560
Db	1501	TTGTGATTTCTTAAAGATGACTATATATTTTCCACTAAAATATTTGTTCTGC	1560

RESULT 5

ABK10034 standard; cDNA; 1674 BP.

XX

XX

AC

AC

21-MAY-2002 (first entry)

XX

Human vascular endothelial growth factor 2 (VEGF-2) cDNA.

DE

XX

Human; vascular endothelial growth factor 2; VEGF-2; gene; ss; pVGI.1;

XX

chronic limb ischaemia; myocardial ischaemia; autoimmune disorder;

KW

allergic reaction; organ rejection; inflammatory condition; arrhythmia;

KW

hyperproliferative disorder; viral infection; bacterial infection;

KW

fungal infection; parasitic infection; cardiovascular disorder; embolism;

KW

heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.

XX

Homo sapiens.

OS

XX

Key

Location/Qualifiers

CDs

12..1271

/*tag= a

/product= "Human VEGF-2"

XX

WO200211769-A1.

XX

14-FEB-2002.

XX

03-AUG-2001; 2001WO-US024658.

XX

04-AUG-2000; 2000US-0223276P.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

Coleman TA;

XX

WPI; 2002-217153/27.

XX

P-PSDB; AAU76812.

XX

Isolated nucleic acid having expression vector construct with vascular

XX

endothelial growth factor-2 insert, useful for treating chronic limb

XX

ischaemia or myocardial ischaemia, autoimmune disorders and allergic

XX

conditions.

XX

Example 9; Fig 1; 241pp; English.

XX

The invention relates to an isolated nucleic acid comprising pVGI.1

XX

expression vector construct containing the vascular endothelial growth

XX

factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host

XX

cell by transducing, transforming or transfecting a host cell with the

XX

DNA and for treating a patient having chronic limb ischaemia or

XX

myocardial ischaemia, or a disease or disorder selected from autoimmune

XX

disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or

XX

conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.

XX

Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),

XX

diseases due to viral, bacterial, fungal or parasitic infection,

XX

cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve

XX

diseases, aneurysms, arterial occlusive disorders and embolism. This

XX

sequence represents cDNA encoding human VEGF-2

XX

Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;

XX

Query Match

100.0%; Score 1674; DB 6; Length 1674;

XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Albert VR, Ruben SM, Wager RE;
 PI WPI; 2003-092991/08.
 DR P-FSDB; ABJ19819.
 XX
 XX New isolated polynucleotide encoding an antibody which inhibits a VEGF-2
 PT polypeptide, useful for diagnosing, treating or preventing diseases
 PT associated with aberrant VEGF-2 expression or function, e.g. cancer or
 PT inflammation.
 XX
 PS Disclosure; Fig 1; 425pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a first
 CC antibody at least 95-100% identical to a second antibody comprising an
 CC amino acid sequence selected from at least one, two or three CDR
 CC region(s) of a VH or VL domain where the first antibody
 CC immunospecifically inhibits a VEGF-2 polypeptide. The isolated
 CC polynucleotide is useful in diagnosing, treating, preventing, prognosing,
 CC ameliorating or monitoring diseases associated with aberrant VEGF-2 or
 CC VEGF-2 receptor expression or lack of VEGF-2 or VEGF-2 receptor function,
 CC such as cancer and other proliferative disorders, cardiovascular
 CC disorders (arrhythmias), cerebrovascular disorders (e.g. cerebral
 CC anoxia), inflammatory diseases, infectious diseases, autoimmune diseases
 CC (e.g. rheumatoid arthritis, Systemic Lupus Erythematosus, allergies),
 CC diabetic retinopathy or psoriasis. The polynucleotide, polypeptide and
 CC antibodies may also be used to stimulate angiogenesis, wound healing, and
 CC promoting vascular tissue repair. The polynucleotide and polypeptide may
 CC also be used for in vitro purposes related to scientific research,
 CC synthesis of DNA and manufacture of DNA vectors, and for the production
 CC of diagnostics and therapeutics to treat human diseases. This
 CC polynucleotide sequence represents a human VEGF-2 DNA of the invention
 XX
 SQ Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 1674; DB 7; Length 1674;
 Best Local Similarity 100.0%; Pred No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	481	TGTCGGCTTACAGATGTGGGGTTCCTGCAATAGTAGAGGGCTGCAGTGCATGAACACCA	540
DB	481	TGTCGGCTTACAGATGTGGGGTTCCTGCAATAGTAGAGGGCTGCAGTGCATGAACACCA	540
QY	541	GCACGAGCTACCTCAGCAAGACGTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
DB	541	GCACGAGCTACCTCAGCAAGACGTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
QY	601	AACAGTAACAATCAGTTTGGCAATCAGTTCCTGCCGATGCATGTCTTAACTGATG	660
DB	601	AACAGTAACAATCAGTTTGGCAATCAGTTCCTGCCGATGCATGTCTTAACTGATG	660
QY	661	TTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGC	720
DB	661	TTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGC	720
QY	721	AGCAGCAACAGACCTGCCCAATTAATCATGTGGAATTAATCATCTGCAGATGCC	780
DB	721	AGCAGCAACAGACCTGCCCAATTAATCATGTGGAATTAATCATCTGCAGATGCC	780
QY	781	TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC	840
DB	781	TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC	840
QY	841	ATGACATCTGTGGCAACAAAGAGCTGATGAAGACCTGTCTAGTGTCTGCGAG	900
DB	841	ATGACATCTGTGGCAACAAAGAGCTGATGAAGACCTGTCTAGTGTCTGCGAG	900
QY	901	CGGGGCTTCGGCTCGCAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT	960
DB	901	CGGGGCTTCGGCTCGCAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT	960
QY	961	GTGTCGTGTAACAAACAACTTCCCGACCAATGTGGGCCCAACCGAGAAATTTGATGAA	1020
DB	961	GTGTCGTGTAACAAACAACTTCCCGACCAATGTGGGCCCAACCGAGAAATTTGATGAA	1020
QY	1021	ACACATGCCAGTGTGTATGTATAAAGAACTCCGCCAGAAATCAACCCCTAAATCCTGAA	1080
DB	1021	ACACATGCCAGTGTGTATGTATAAAGAACTCCGCCAGAAATCAACCCCTAAATCCTGAA	1080
QY	1081	ATGTGCTGTGAATGTACAGAACTCCAGAAATGCTTGTATAAAGAAAGAGTTC	1140
DB	1081	ATGTGCTGTGAATGTACAGAACTCCAGAAATGCTTGTATAAAGAAAGAGTTC	1140
QY	1141	ACCACCAACATGACAGCTGTGTACAGACGCCCATGTAGCAACCGCCAGAGCTTGTGAGC	1200
DB	1141	ACCACCAACATGACAGCTGTGTACAGACGCCCATGTAGCAACCGCCAGAGCTTGTGAGC	1200
QY	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATGGCAAGACAC	1260
DB	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATGGCAAGACAC	1260
QY	1261	AAATGAGCTTAAGATTTGATGTTTCCAGATTCATCGATTTTCTATTATGAAATCTGT	1320
DB	1261	AAATGAGCTTAAGATTTGATGTTTCCAGATTCATCGATTTTCTATTATGAAATCTGT	1320
QY	1321	TGCCAGTGTAACTGTCTGTGAACAGAGACCCCTGTGGGTCCATGCTTAAAGACA	1380
DB	1321	TGCCAGTGTAACTGTCTGTGAACAGAGACCCCTGTGGGTCCATGCTTAAAGACA	1380
QY	1381	AAAGTCTGTCTTCTCGTAAACCATGTGATAACTTTACAGAAATGGAGCTCATCTG	1440
DB	1381	AAAGTCTGTCTTCTCGTAAACCATGTGATAACTTTACAGAAATGGAGCTCATCTG	1440
QY	1441	CAAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCTC	1500
DB	1441	CAAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCTC	1500
QY	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTTCCACTAAAATATTTTCTG	1560
DB	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTTCCACTAAAATATTTTCTG	1560
QY	1561	ATTCAATTTTATAGCAACAAATGTTGGTAAACTCAGTGTATCAATATTTTATATCAT	1620

Db 1561 ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
 Qy 1621 GCAAAATATGTTTAAATATAATGAAATGTTATTTATATAAAAAAATTTTAAAAA 1674
 Db 1621 GCAAAATATGTTTAAATATAATGAAATGTTATTTATATAAAAAAATTTTAAAAA 1674

RESULT 7

ABQ76966 standard; cDNA; 1674 BP.

XX AC ABQ76966;

XX DT 03-APR-2003 (first entry)

XX DE Human VEGF-2 cDNA SEQ ID 1.

XX KW Human; VH domain; VL domain; vascular endothelial growth factor; VEGF-2;
 KW cytotatic; cardiac; cardiovascular; antiinflammatory; antirheumatic;
 KW antiarthritic; antidiabetic; ophthalmological; antiallergic; vulnerary;
 KW immunosuppressive; dermatological; antipsoriatic; proliferative disorder;
 KW cancer; cardiovascular disorder; arrhythmia; cerebrovascular disorder;
 KW cerebral anoxia; inflammatory disease; infectious disease; angiogenesis;
 KW autoimmune disease; Systemic Lupus Erythematosus; wound healing;
 KW vascular tissue repair; Gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO200283850-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011405.

XX PR 13-APR-2001; 2001US-0283408P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Albert VR, Ruben SM, Wager RE;

XX PT WPI; 2003-093008/08.

XX PS P-PSDB; ABG73702.

PT New isolated polynucleotide encoding an antibody which inhibits vascular
 PT endothelial growth factor (VEGF)-2 polypeptide, useful for diagnosing or
 PT treating diseases associated with aberrant VEGF-2 expression or function,
 PT e.g. cancer.

XX PS Disclosure; Fig 1A-B; 344pp; English.

XX This invention describes a novel isolated polynucleotide encoding a first
 CC antibody comprising an amino acid sequence selected from at least one,
 CC two or three complementarity determining (CDR) region(s) of a VH and/or
 CC VL domain of a second antibody that immunospecifically binds to a
 CC vascular endothelial growth factor (VEGF)-2 polypeptide. The products of
 CC the invention have cytostatic, cardiovascular, antirheumatic, vulnerary,
 CC antiinflammatory, antiarthritic, antidiabetic, ophthalmological, and
 CC cardiac, antiallergic, immunosuppressive, dermatological and
 CC antipsoriatic activity. The polynucleotide is useful in diagnosing,
 CC treating, preventing, prognosing, ameliorating or monitoring diseases
 CC associated with aberrant VEGF-2 or VEGF-2 receptor expression or lack of
 CC VEGF-2 or VEGF-2 receptor function, such as cancer and other
 CC proliferative disorders, cardiovascular disorders (arrhythmias),
 CC cerebrovascular disorders (e.g. cerebral anoxia), inflammatory diseases,
 CC infectious diseases, autoimmune diseases (e.g. rheumatoid arthritis,
 CC Systemic Lupus Erythematosus, allergies), diabetic retinopathy or
 CC psoriasis. The polynucleotide, polypeptide and antibodies described in
 CC the invention may also be used to stimulate angiogenesis, wound healing,
 CC and promoting vascular tissue repair. The polynucleotide and polypeptide
 CC may also be used for in vitro purposes related to scientific research,
 CC synthesis of DNA and manufacture of DNA vectors and for the production of
 CC diagnostics and therapeutics e.g. gene therapy and in vaccines to treat

CC human diseases. This sequence encodes a polypeptide described in the
 CC disclosure of the invention
 XX
 SQ Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1674; DB 7; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCCTTCCACCATGCACTCGCTGGGCTTCTTCTCTGTGGGCTTCTTCTCTGTGGGCTG 60
 DB 1 GTCCTTCCACCATGCACTCGCTGGGCTTCTTCTCTGTGGGCTTCTTCTCTGTGGGCTG 60
 QY 61 CGCTGCTCCCGGCTCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
 DB 61 CGCTGCTCCCGGCTCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
 QY 121 ACCTCTCGGACGGGAGCCCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCTGG 180
 DB 121 ACCTCTCGGACGGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCTGG 180
 QY 181 AGGAGCAGTTACGCTCTGTCTCCAGTGTAGTGAATCATGACTGTACTTCTACCCAGAT 240
 DB 181 AGGAGCAGTTACGCTCTGTCTCCAGTGTAGTGAATCATGACTGTACTTCTACCCAGAT 240
 QY 241 ATTGGAAAAATGTACAAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCA 300
 DB 241 ATTGGAAAAATGTACAAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCA 300
 QY 301 CCAACCTCAACTCAAGGACAG 360
 DB 301 CCAACCTCAACTCAAGGACAG 360
 QY 361 AGATCTTGAAGATTTGATATAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 DB 361 AGATCTTGAAGATTTGATATAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 421 GTATAGATGTGGGAG 480
 DB 421 GTATAGATGTGGGAG 480
 QY 481 TGTCCGCTACAGATGTGGGAG 540
 DB 481 TGTCCGCTACAGATGTGGGAG 540
 QY 541 GCACGAGCTACCTCAGCAAG 600
 DB 541 GCACGAGCTACCTCAGCAAG 600
 QY 601 AACCGAGTAACTCAAGTGTGGCAATCACTTCTGCGAGTGCATGTCTTAACCTGAGATG 660
 DB 601 AACCGAGTAACTCAAGTGTGGCAATCACTTCTGCGAGTGCATGTCTTAACCTGAGATG 660
 QY 661 TTTACAGACAAAGTTCACTTCAATTTAGACGTTCCCTGCCAGCAACATCAACAGATGTC 720
 DB 661 TTTACAGACAAAGTTCACTTCAATTTAGACGTTCCCTGCCAGCAACATCAACAGATGTC 720
 QY 721 AGGAGAGCAAG 780
 DB 721 AGGAGAGCAAG 780
 QY 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATGCC 840
 DB 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATGCC 840
 QY 841 ATGACATCTGTGACCAAAACAG 900
 DB 841 ATGACATCTGTGACCAAAACAG 900
 QY 901 CGGAGCTTGGGCTCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 DB 901 CGGAGCTTGGGCTCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

QY 961 GTGCTCTTAAACAACTCTTCCCGAGCAATGTGGGSCCAACCGAGAAATTTGATGAAA 1020
 DB 961 GTGCTCTTAAACAACTCTTCCCGAGCAATGTGGGSCCAACCGAGAAATTTGATGAAA 1020
 QY 1021 ACATATCCAGTGTGTATGTAAAGAACTGCGCCAGAGAAATCAACCCCTTAATCTCGAA 1080
 DB 1021 ACATATCCAGTGTGTATGTAAAGAACTGCGCCAGAGAAATCAACCCCTTAATCTCGAA 1080
 QY 1081 AATGTGCTGTATGTACAGAAAGTCCAGAAATGCTTGTAAAGAAAGAAAGTCTCC 1140
 DB 1081 AATGTGCTGTATGTACAGAAAGTCCAGAAATGCTTGTAAAGAAAGAAAGTCTCC 1140
 QY 1141 ACCACCAACATGAGCTGTACAGAGCGCCATGTACGAAACGCGCAGAGCTTTGTGAGC 1200
 DB 1141 ACCACCAACATGAGCTGTACAGAGCGCCATGTACGAAACGCGCAGAGCTTTGTGAGC 1200
 QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCCTTCATATTGGCAAGACAC 1260
 DB 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCCTTCATATTGGCAAGACAC 1260
 QY 1261 AATGAGCTAAGATTTGTATCTGTTTCCAGTTTCATGATTTCTATATTGGAAGAACTGTGT 1320
 DB 1261 AATGAGCTAAGATTTGTATCTGTTTCCAGTTTCATGATTTCTATATTGGAAGAACTGTGT 1320
 QY 1321 TGCCACAGTGAAGTGTGTGTAACAGAGAGACCCCTTGTGGTCCATGCTAACAAAGACA 1380
 DB 1321 TGCCACAGTGAAGTGTGTGTAACAGAGAGACCCCTTGTGGTCCATGCTAACAAAGACA 1380
 QY 1381 AATGCTGTCTTCCGACCATGTGATACTTTACAGAAATGAGTGGAGCTCATCTG 1440
 DB 1381 AATGCTGTCTTCCGACCATGTGATACTTTACAGAAATGAGTGGAGCTCATCTG 1440
 QY 1441 CAAAGGCTCTTGTAAAGAGTGTGTTTCCGCAATGACCAACAGCAGCAAGTTTCTC 1500
 DB 1441 CAAAGGCTCTTGTAAAGAGTGTGTTTCCGCAATGACCAACAGCAGCAAGTTTCTC 1500
 QY 1501 TTGATGTTCTTAAAGAGTGTATATAATTTTCCACTAAAATATTGTTCTGC 1560
 DB 1501 TTGATGTTCTTAAAGAGTGTATATAATTTTCCACTAAAATATTGTTCTGC 1560
 QY 1561 ATTCAATTTTATAGCAACAAATGTTGTAATCTCACTGTGATCAATTTTATATCAT 1620
 DB 1561 ATTCAATTTTATAGCAACAAATGTTGTAATCTCACTGTGATCAATTTTATATCAT 1620
 QY 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTATAAATAAATAAATAAATAA 1674
 DB 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTATAAATAAATAAATAAATAA 1674

RESULT 8

AAD49522
 ID AAD49522 standard; cDNA; 1674 BP.
 AC AAD49522;
 XX 24-MAR-2003 (first entry)
 XX Human vascular endothelial growth factor (VEGF-2) cDNA.
 DE Human; vascular endothelial growth factor; VEGF-2; inflammatory disease;
 KW proliferative disorder; tumour; breast; cancer; brain; prostate; colon;
 KW lymphangioma; infection; Kaposi's sarcoma; psoriasis; immunosuppressive;
 KW rheumatoid arthritis; diabetic retinopathy; gene therapy; antimicrobial;
 KW cytosolic; ophthalmological; autoimmune disease; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 12..1271
 FT /*tag= a
 FT /*product= "Human VEGF-2 protein"
 FT 12..80
 FT /*tag= b

mat_peptide 81..1268
 /*tag= c
 /product= "Human mature VEGF-2 protein"
 WO200283849-A2.
 24-OCT-2002.
 12-APR-2002; 2002WO-US011404.
 13-APR-2001; 2001US-0283391P.
 07-SEP-2001; 2001US-0317600P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Albert VR, Ruben SM, Wager RE;
 WPI, 2003-093007/08.
 P-PSDB; AAE32091.
 New vascular endothelial growth factor (VEGF)- 2 antibodies, for
 treating, preventing or ameliorating a disease or disorder, such as
 inflammatory diseases, proliferative disorders, autoimmune disorders or
 diabetic retinopathy.
 Example 9; Page 364-366; 399pp; English.
 The invention relates to vascular endothelial growth factor (VEGF)-2
 antibodies. VEGF-2 antibodies are useful for treating, preventing or
 ameliorating a disease or disorder, such as inflammatory diseases or
 disorders, proliferative disorders, tumours, tumour metastasis, breast
 cancer, brain cancer, prostate cancer, colon cancer, lymphangioma, an
 infectious disease, Kaposi's sarcoma, an autoimmune disease, rheumatoid
 arthritis, psoriasis, diabetic retinopathy, a disease or disorder
 associated with aberrant VEGF-2 (receptor) expression, or a disease or
 disorder associated with the lack of VEGF-2 (receptor) function. The
 antibody is also useful for detecting, diagnosing, prognosing, or
 monitoring cancers and other hyperproliferative disorders. VEGF-2 is also
 used in gene therapy. The present sequence is human VEGF-2 cDNA
 Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1674; DB 7; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCTTCCACATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCGCCGCTG 60
 DB 1 GTCTTCCACATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCGCCGCTG 60
 QY 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCCGCGCGCGCGCGCTTCGAGTCCGGACTCG 120
 DB 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCCGCGCGCGCGCTTCGAGTCCGGACTCG 120
 QY 121 ACCTCTCGAGCGGAGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
 DB 121 ACCTCTCGAGCGGAGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
 QY 181 AGGAGCAGTTACGCTGTGTCTGCTAGTGTAGTGAATCTGCTGCTACTCTACCCAGAT 240
 DB 181 AGGAGCAGTTACGCTGTGTCTGCTAGTGTAGTGAATCTGCTGCTACTCTACCCAGAT 240
 QY 241 ATTGGAAAAATGTACAAGTGTCTAGCTTAAGGAAAGAGGCTGGCAACATACAGAAACAGG 300
 DB 241 ATTGGAAAAATGTACAAGTGTCTAGCTTAAGGAAAGAGGCTGGCAACATACAGAAACAGG 300
 QY 301 CCACTCTCACTCAAGCAGAGAGAGACTATAAATTTGCTGCGAGCATTATATACAG 360
 DB 301 CCACTCTCACTCAAGCAGAGAGAGACTATAAATTTGCTGCGAGCATTATATACAG 360
 QY 361 AGATCTTGAAGAACTATTGATAATGAGTGGAGAAAGAGACTCAATGATGCCAGGAGGTGT 420
 DB 361 AGATCTTGAAGAACTATTGATAATGAGTGGAGAAAGAGACTCAATGATGCCAGGAGGTGT 420

QY	421	GTATAGATGTGGGAAAGAGTTTGGAGTGGCGCAAAACACCTTCTTTAAACCTCCATGTG	480
DB	421		
QY	421	GTATAGATGTGGGAAGAGTTTGGAGTCGCGCAAAACACCTTCTTTAAACCTCCATGTG	480
DB	421		
QY	481	TGTCGGTCTTACAGATGTGGGGGTGTGCAATAGTGGGGGTGCAGTGCATGAACACCA	540
DB	481		
QY	481	TGTCGGTCTTACAGATGTGGGGGTGTGCGAATAGTGGGGGTGCAGTGCATGAACACCA	540
DB	481		
QY	541	GCACGAGCTACTCAGACAGACGTTATTGAAATATACAGTGCCCTCTCTCAAGGCCCCA	600
DB	541		
QY	541	GCACGAGCTACTCAGACAGACGTTATTGAAATATACAGTGCCCTCTCTCAAGGCCCCA	600
DB	541		
QY	601	TAACAGTAACAAATCAGTTTTTGGCCAATCACACTTCTGCGGATGCATGTCTAAATCGATG	660
DB	601		
QY	601	TAACAGTAACAAATCAGTTTTTGGCCAATCACACTTCTGCGGATGCATGTCTAAATCGATG	660
DB	601		
QY	661	TTTACAGACAGTTTCAATTCATTTATAGCGTCCCTGCCAGACACTACACAGTGTC	720
DB	661		
QY	661	TTTACAGACAGTTTCAATTCATTTATAGCGTCCCTGCCAGACACTACACAGTGTC	720
DB	661		
QY	721	AGGCAGCAACAAAGACCTGCCCAACCAATTACATGTGGAATTAATCACATCTGCAGATGCC	780
DB	721		
QY	721	AGGCAGCAACAAAGACCTGCCCAACCAATTACATGTGGAATTAATCACATCTGCAGATGCC	780
DB	721		
QY	781	TGGCTCAGGAGATTTTATGTTTTTCTCGAGTCTGAGATGACATCAACAGATGATTC	840
DB	781		
QY	781	TGGCTCAGGAGATTTTATGTTTTTCTCGAGTCTGAGATGACATCAACAGATGATTC	840
DB	781		
QY	841	ATGACATCTGTGGACCAAAACAGAGCTGGATGAGAGACCTGTCACTGTCTGCAGAG	900
DB	841		
QY	841	ATGACATCTGTGGACCAAAACAGAGCTGGATGAGAGACCTGTCACTGTCTGCAGAG	900
DB	841		
QY	901	CGGGCTTCGGCTCGCACCTGTGAGACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT	960
DB	901		
QY	901	CGGGCTTCGGCTCGCACCTGTGAGACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT	960
DB	901		
QY	961	GTGCTGTGTAACAACTCTTCCCGACCAATGTGGGCCCAACCGAGAAATTGATGAAA	1020
DB	961		
QY	961	GTGCTGTGTAACAACTCTTCCCGACCAATGTGGGCCCAACCGAGAAATTGATGAAA	1020
DB	961		
QY	1021	ACACATGCCAGTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGGAA	1080
DB	1021		
QY	1021	ACACATGCCAGTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGGAA	1080
DB	1021		
QY	1081	AAATGCTGCTGTGATGTACAGAAAGTCCACAGAAATCTGTGTTAAAGGAAGAGTTCC	1140
DB	1081		
QY	1081	AAATGCTGCTGTGATGTACAGAAAGTCCACAGAAATCTGTGTTAAAGGAAGAGTTCC	1140
DB	1081		
QY	1141	ACCACCAACATGCAGCTGTTTACAGAGCGCCATGTACGAACCGCCAGAGGCTTTGTGAGC	1200
DB	1141		
QY	1141	ACCACCAACATGCAGCTGTTTACAGAGCGCCATGTACGAACCGCCAGAGGCTTTGTGAGC	1200
DB	1141		
QY	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTCCCTTCATATTATGGCAAGACAC	1260
DB	1201		
QY	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTCCCTTCATATTATGGCAAGACAC	1260
DB	1201		
QY	1261	AAATGAGCTAAGATTGTAAGTTTCCAGTTTCAATGATTTTCTATTATGGAATCTGT	1320
DB	1261		
QY	1261	AAATGAGCTAAGATTGTAAGTTTCCAGTTTCAATGATTTTCTATTATGGAATCTGT	1320
DB	1261		
QY	1321	TGCCAGTGAAGCTGTCTGTGAACAGAGAGACCTTTGTGGTCCATGTAAACAAAGACA	1380
DB	1321		
QY	1321	TGCCAGTGAAGCTGTCTGTGAACAGAGAGACCTTTGTGGTCCATGTAAACAAAGACA	1380
DB	1321		
QY	1381	AAAGTCTGCTTTCCTGAAACCATGTGATTAACCTTTACGAAATGAGCTCATCTG	1440
DB	1381		
QY	1381	AAAGTCTGCTTTCCTGAAACCATGTGATTAACCTTTACGAAATGAGCTCATCTG	1440
DB	1381		
QY	1441	CAAAAGGCTCTTGTAAAGATGCTTTTCTGCCAATGACCAACAGCAAGATTTCTCTC	1500
DB	1441		
QY	1441	CAAAAGGCTCTTGTAAAGATGCTTTTCTGCCAATGACCAACAGCAAGATTTCTCTC	1500
DB	1441		

Qy	1501	TTGTGGATTCCTTTAAAGAAATGACIATATATAATTTATTTCCACTAAAAATATATGTTTCTGC	1560
Db	1501	TTGTGGATTCCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATGTTTCTGC	1560
Qy	1561	ATTCAATTTTATAGCAACACAAATTCGTAAAACTCCTGTGATCAATATTTTATATCAT	1620
Db	1561	ATTCAATTTTATAGCAACACAAATTCGTAAAACTCCTGTGATCAATATTTTATATCAT	1620
Qy	1621	GCAAAATATGTTTAAATAAAAATGAAATATGTAATTTATAAAAAATAAAAAA	1674
Db	1621	GCAAAATATGTTTAAATAAAAATGAAATATGTAATTTATAAAAAATAAAAAA	1674

6. 4. 11. 2019

RESOL 9
AAT51371
ID AAT51371 standard: cDNA: 1674 BP.

AA
AC
AA51371.

23-APR-1997 (first entry)

XX DE Human vascular endothelial growth factor 2 cDNA.

XX	
KW	Vascular endothelial growth factor 2; VEGF2; angiogenesis;
KW	endothelialisation; coronary bypass surgery; vascular graft surgery,
KW	agonist; antagonist; gene therapy; diagnosis; ds.

XX Homo sapiens.

XX		
FH	Key	Location/Qualifiers
FT	CDS	12..1271
FT		/tag= a
FT	sig_peptide	12..149
FT		/tag= b
FT	mat_peptide	150..1268
FT		/tag= c
FT	primer_bind	complement(150..166)
FT		/tag= d
FT	'	/note= "5' primer binding site"
FT	primer_bind	1254..1268
FT		/tag= e
FT		/note= "3' primer binding site"
FT	primer_bind	1254..1268
FT		/tag= f
FT		/note= "3' primer binding site"
FT		

XX PN W00639515-A1

XX
PD
12-DEC-1996XX
PF
06-JTN-1996;XX
PB
06-TTN-1995- 95TS-00465968XX
PA
(HTMA-) HTMAN GENOME SGT TNC

XX
PT
Rosen CA
Hill J
Cao L:

XX
NR
WPT: 1997-043137/04

DR P-PSDB; AAW11478.
yy

PT DNA encoding human angiotensinogen or e

XX
PS
Claim 1: Fig 1: 74pp: English

CC A cDNA clone (AAT51371) codes

CC
FACCTOF 2 (VEGF2) (AAWI4/8);
CC
VEGF/PDGF family that is capable

9. VEGF2 polynucleotides may also be obtained from adult heart or several breast cancer cell lines. VEGF2 nucleic acids can be used in the production of recombinant VEGF2, as probes to detect mutations in the VEGF2, and in

CC Gene therapy to treat patients in need of VEGF2. Antisense sequences can
 CC be used as VEGF2 antagonists e.g. to inhibit growth of tumours
 XX
 SQ Sequence 1674 BP; 502 A; 384 C; 375 G; 413 T; 0 U; 0 Other;

Query Match 99.68; Score 1667.6; DB 2; Length 1674;
 Best Local Similarity 99.88; Pred. No. 0;
 Matches 1670; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCTCGTGGGCTTCTCTGTGGGCTGTTCTCTGCTCCGCGCTG 60
 Db 1 GTCTTCCACCATGCTCGTGGGCTTCTCTGTGGGCTGTTCTCTGCTCCGCGCTG 60

Qy 61 CGCTGCTCCGGGTCCTCGGAGCGCCGCGCGCGCGCGCGCTTCGAGTCGGACTCG 120
 Db 61 CGCTGCTCCGGGTCCTCGGAGCGCCGCGCGCGCGCGCGCTTCGAGTCGGACTCG 120

Qy 121 ACCTCTCGGAGCGCGGAGCCCGACCGCGGCGAGGCCAGGCTTATGCAAGCAAGATCTG 180
 Db 121 ACCTCTCGGAGCGCGGAGCCCGACCGCGGCGAGGCCAGGCTTATGCAAGCAAGATCTG 180

Qy 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAACTCATGCTGTACTCTACCCAGAT 240
 Db 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAACTCATGCTGTACTCTACCCAGAT 240

Qy 241 ATTGGAATAATGTAAGTGTGAGTAAAGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
 Db 241 ATTGGAATAATGTAAGTGTGAGTAAAGAAAGAGGCTGGCAACATAACAGAGAACAGG 300

Qy 301 CCAACCTCACTAAGGACAGAGAGACTATAAATTTGTCGAGCAACATTAATACAG 360
 Db 301 CCAACCTCACTAAGGACAGAGAGACTATAAATTTGTCGAGCAACATTAATACAG 360

Qy 361 AGATCTTGAAAGATTGTAATGAGTGGAGAGAGCTCAATGATGCCACGGAGGTGT 420
 Db 361 AGATCTTGAAAGATTGTAATGAGTGGAGAGAGCTCAATGATGCCACGGAGGTGT 420

Qy 421 GTATAGATGTGGGAGAGGTTGGAGTTCGCGCAAAACACCTTTCTTAAACCTCCATGTG 480
 Db 421 GTATAGATGTGGGAGAGGTTGGAGTTCGCGCAAAACACCTTTCTTAAACCTCCATGTG 480

Qy 481 TGTCCGCTCAGATGTGGGGTGTCTGCAATAGTGGAGGGCTGCAGTGCATGACACCA 540
 Db 481 TGTCCGCTCAGATGTGGGGTGTCTGCAATAGTGGAGGGCTGCAGTGCATGACACCA 540

Qy 541 GCAGGAGTACTCTCAGCAAGACGTTATTGAAATTAACAGTGCCTCTCTCTCAAGGCCCCA 600
 Db 541 GCAGGAGTACTCTCAGCAAGACGTTATTGAAATTAACAGTGCCTCTCTCTCAAGGCCCCA 600

Qy 601 AACGAGTAAATCAGATTTGCGCAATCAGCTTCTGCGGATGATGCTTAAACTGGATG 660
 Db 601 AACGAGTAAATCAGATTTGCGCAATCAGCTTCTGCGGATGATGCTTAAACTGGATG 660

Qy 661 TTTACAGCAAGTTTCAATTCATTTAGAGCTTCCCTGCGGAGCAACACTACCAAGTGC 720
 Db 661 TTTACAGCAAGTTTCAATTTAGAGCTTCCCTGCGGAGCAACACTACCAAGTGC 720

Qy 721 AGGAGCGGAAACAGACTGCCCCACCAATTACATGTGGAATAATACATCTGCAGATGCC 780
 Db 721 AGGAGCGGAAACAGACTGCCCCACCAATTACATGTGGAATAATACATCTGCAGATGCC 780

Qy 781 TGGCTCAGGAGATTTATGTTTCTCCGATGCTGGAGTGAATCAACAGATGGATTC 840
 Db 781 TGGCTCAGGAGATTTATGTTTCTCCGATGCTGGAGTGAATCAACAGATGGATTC 840

Qy 841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGAGTGTCTGCGAGAG 900
 Db 841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGAGTGTCTGCGAGAG 900

Qy 901 CGGGCTTCGGCTTCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
 Db 901 CGGGCTTCGGCTTCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960

Qy 961 GTGTCTGTAAAAACAACAACTCTTCCAGCCCAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
 Db 961 GTGTCTGTAAAAACAACAACTCTTCCAGCCCAATGTGGGCGCAACCGCAATTTGATGAAA 1020

Qy 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGCAGAAATCAACCCCTTAATCCTGGAA 1080
 Db 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGCAGAAATCAACCCCTTAATCCTGGAA 1080

Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAAAGAAAGAGTTCC 1140
 Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAAAGAAAGAGTTCC 1140

Qy 1141 ACCACCAAAATGAGCTGTTTACAGAGCGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1200
 Db 1141 ACCACCAAAATGAGCTGTTTACAGAGCGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1200

Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCTATATTGGAAGAACCTG 1260
 Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCTATATTGGAAGAACCTG 1260

Qy 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTCATCGATTTTCTATTATGGAAGAACCTG 1320
 Db 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTCATCGATTTTCTATTATGGAAGAACCTG 1320

Qy 1321 TGCCACAGTACAGTCTGTGAAACAGAGAGACCTTGTGGTCCATGCTAACAAAGACA 1380
 Db 1321 TGCCACAGTACAGTCTGTGAAACAGAGAGACCTTGTGGTCCATGCTAACAAAGACA 1380

Qy 1381 AAAGTCTGCTTCTGTAACCATGTGGTAACTTTTACAGAAATGAGCTGGAGCTCATCTG 1440
 Db 1381 AAAGTCTGCTTCTGTAACCATGTGGTAACTTTTACAGAAATGAGCTGGAGCTCATCTG 1440

Qy 1441 CAAAAGGCTCTCTGTAAAGAGCTGTTTCTGCCAATGACCAAAACAGCAAGATTTTCTC 1500
 Db 1441 CAAAAGGCTCTCTGTAAAGAGCTGTTTCTGCCAATGACCAAAACAGCAAGATTTTCTC 1500

Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
 Db 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560

Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
 Db 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620

Qy 1621 GCAAAATATGTTTAAATATAATGAAATGTTATTTATATAAAAAA 1674
 Db 1621 GCAAAATATGTTTAAATATAATGAAATGTTATTTATAAAAAA 1674

RESULT 10
 AAT59929
 ID AAT59929 standard; cDNA; 2031 BP.
 XX
 AC AAT59929;
 XX
 DT 05-JUN-1997 (first entry)
 XX
 DE Human vascular endothelial growth factor-related protein VRP cDNA.
 KW Vascular endothelial growth factor-related protein; VRP; VEGF;
 KW receptor protein tyrosine kinase; Flt4; signal transduction;
 KW wound healing; vulvar; rheumatoid arthritis; Kaposi's sarcoma;
 KW therapy; diagnosis; angiogenesis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 372..1631
 FT /*tag= a
 FT /transl_except= pos:711..713_aa:Thr
 FT sig_peptide
 FT 372..431
 FT /*tag= b
 FT mat_peptide
 FT 432..1628

FT	XX	PN	XX	PD	XX	PF	XX	PR	XX	PS	XX	PT	XX	DR	XX	DR	XX	PT	XX	PT	XX	PS	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX	CC	XX
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Db 1921 ATTCAATTTTATAGCAACAACTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1980

QY 1621 GCAAAATATGTTAAATTAATGAAATGTATTATTAATAAAAAA 1671

Db 1981 GCAAAATATGTTAAATTAATGAAATGTATTATTAATAAAAAA 2031

RESULT 11

AA84276

ID AAT84276 standard; cDNA; 1997 BP.

XX AAT84276;

XX

DT 10-NOV-1997 (first entry)

XX Human Flt4 receptor tyrosine kinase ligand VEGF-C cDNA.

DE

XX VEGF-C; Flt4; receptor tyrosine kinase; VEGFR-3; human;

KW vascular endothelial growth factor receptor-3; ligand; angiogenesis;

KW wound healing; lymph vessel; lymphangioma; cancer; metastasis; therapy;

KW diagnosis; plasmid pFLT-4-L; ss.

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT CDS 352..1611

FT /*tag= a

FT sig_peptide 352..657

FT /*tag= b

FT mat_peptide 658..1608

FT /*tag= c

XX

PN WO9705250-A2.

XX

PD 13-FEB-1997.

XX

XX 01-AUG-1996; 96WO-FI000427.

XX

PR 01-AUG-1995; 95US-00510133.

PR 12-JAN-1996; 96US-00585895.

PR 14-FEB-1996; 96US-00601132.

PR 28-JUN-1996; 96US-00671573.

XX

XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.

XX

PI Alitalo K, Joukov V;

XX

DR WPI; 1997-145688/13.

DR P-PSDB; AAW00932.

XX

XX Flt4 receptor tyrosine kinase ligand and related nucleic acid - used to

PT modulate growth of endothelial cells and for diagnosis of endothelial

PT cell diseases.

XX

XX Claim 25; Page 110-112; 183pp; English.

XX

XX This cDNA clone codes for human VEGF-C (AAW00932), a novel ligand that

CC binds specifically to human Flt4 receptor tyrosine kinase (VEGFR-3),

CC stimulating phosphorylation of the receptor, and which can be used to

CC modulate growth of endothelial cells and for diagnosis of endothelial

CC cell diseases. The clone was isolated from a PC-3 prostatic

CC adenocarcinoma cell (ATCC CRL 1435) cDNA library by PCR amplification and

CC has been cloned into pcDNA1 to form plasmid pFLT4-L (ATCC 97231). The

CC human VEGF-C gene is located on chromosome 4q34 and contains 6 introns

CC (see A184303-14). The insert from the human clone was used to isolate

CC mouse (see A184277) and quail (see AAT84300) VEGF-C sequences. It can

CC also be used to produce recombinant VEGF-C polypeptides in host cells and

CC to provide probes for screening endothelial cell diseases. Antisense

CC molecules that block the receptor can be used to control endothelial cell

CC proliferation, e.g. lymphangioma or metastatic cancer. The promoter

XX sequence (AAT84278) is also claimed

XX

XX Sequence 1997 BP; 526 A; 527 C; 490 G; 454 T; 0 U; 0 Other;

SQ

Query Match 98.7%; Score 1651.8; DB 2; Length 1997;

Best local similarity 99.9%; Pred. No. 0;

Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCTTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTGCGCGCTG 60

Db 341 GTCTTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTGCGCGCTG 400

QY 61 CGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

Db 401 CGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 460

QY 121 ACCTCTCGAGCGCGGAGCCGAGCGCGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

Db 461 ACCTCTCGAGCGCGGAGCCGAGCGCGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 520

QY 181 AGGAGCAGTTACCGTCTGTGTCTCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAT 240

Db 521 AGGAGCAGTTACCGTCTGTGTCTCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAT 580

QY 241 ATTGGAAATGTACAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 300

Db 581 ATTGGAAATGTACAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 640

QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACATTAATAACAG 360

Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACATTAATAACAG 700

QY 361 AGATCTTGAAGTATTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420

Db 701 AGATCTTGAAGTATTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 760

QY 421 GTATAGATGTGGGAGGAGTTTGGAGTGCAGCAACAAACACCTTCTTTAAACCTCCATGTG 480

Db 761 GTATAGATGTGGGAGGAGTTTGGAGTGCAGCAACAAACACCTTCTTTAAACCTCCATGTG 820

QY 481 TGTCCGTCTACAGATGTGGGGTGTGCTGCAATAGTGTGAGGGGTGAGTGCATGAACACCA 540

Db 821 TGTCCGTCTACAGATGTGGGGTGTGCTGCAATAGTGTGAGGGGTGAGTGCATGAACACCA 880

QY 541 GCACGAGCTACCTCAGCAGAGGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600

Db 881 GCACGAGCTACCTCAGCAGAGGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 940

QY 601 AACCAATCAATCACTGCTTCCATCACTTCTCTGCGGATCATGTCTAACTGGATG 660

Db 941 AACCAATCAATCACTGCTTCCATCACTTCTCTGCGGATCATGTCTAACTGGATG 1000

QY 661 TTACAGACAGTTTCATTCATTTAGAGTTCCTGCGAGGAGTTCCTGCGAGCAACACTACCAAGTGC 720

Db 1001 TTACAGACAGTTTCATTCATTTAGAGTTCCTGCGAGGAGTTCCTGCGAGCAACACTACCAAGTGC 1060

QY 721 AGGCAGCAACAGACCTGCCCCACCAATTACATGTGGAATAATCACATCTGCAGATGCC 780

Db 1061 AGGCAGCAACAGACCTGCCCCACCAATTACATGTGGAATAATCACATCTGCAGATGCC 1120

QY 781 TGGCTCAGGAAGATTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGATTC 840

Db 1121 TGGCTCAGGAAGATTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGATTC 1180

QY 841 ATGACATCTGTGGACCAACAGAGGCTGGATGAGAGACCTGTGCTGTCTGCAGAG 900

Db 1181 ATGACATCTGTGGACCAACAGAGGCTGGATGAGAGACCTGTGCTGTCTGCAGAG 1240

QY 901 CGGGGCTTCGGCTCCAGCTGTGTGACCCCAAGAACTAGACAGAAACTCATGCCAGT 960

Db 1241 CGGGGCTTCGGCTCCAGCTGTGTGACCCCAAGAACTAGACAGAAACTCATGCCAGT 1300

QY 961 GTGTCTGTAATAACAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020

Db 1301 GTGTCTGTAATAACAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1360

1021 ACACATGCGCTGTATGTAAAGACCTGCCCGGAGAAATCAACCCCTAAATCCTGGAA 1080
 1361 ACACATGCGCTGTATGTAAAGACCTGCCCGGAGAAATCAACCCCTAAATCCTGGAA 1420
 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAAAGTTCC 1140
 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAAAGTTCC 1480
 1141 ACCACCAACATGCTGTACAGAGCGCATGTACAGAGCGCATGTACAGAGCGCATGTGAGC 1200
 1481 ACCACCAACATGCTGTACAGAGCGCATGTACAGAGCGCATGTACAGAGCGCATGTGAGC 1540
 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATTTGGCAAGACCCAC 1260
 1541 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATTTGGCAAGACCCAC 1600
 1261 AATGAGCTAAGATGTGCTTTTCCAGTTCAATGATTTTCTATTTATGGAAGAACTGTGT 1320
 1601 AATGAGCTAAGATGTGCTTTTCCAGTTCAATGATTTTCTATTTATGGAAGAACTGTGT 1660
 1321 TGCCACAGTAACTGCTGTGAACAGAGAGCCCTTGTGGTCCATGCTTACAAAGACA 1380
 1661 TGCCACAGTAACTGCTGTGAACAGAGAGCCCTTGTGGTCCATGCTTACAAAGACA 1720
 1381 AAGTCTGTCTTCTGAACTGTGTGAATTAATTTACAGAAATGGAGCTCATCTG 1440
 1721 AAGTCTGTCTTCTGAACTGTGTGAATTAATTTACAGAAATGGAGCTCATCTG 1780
 1441 CAAAAGCCCTCTGTGAAGACTGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTTC 1500
 1781 CAAAAGCCCTCTGTGAAGACTGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTTC 1840
 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTATTTCCACTTAAATAATTTTCTGC 1560
 1841 TTGTGATTTCTTTAAAGAGTACTATATAATTTATTTCCACTTAAATAATTTTCTGC 1900
 1561 ATTCTATTTTATAGCAACAAATTTGGTAAATCTACTGTGATCAATATTTTATATCAT 1620
 1901 ATTCTATTTTATAGCAACAAATTTGGTAAATCTACTGTGATCAATATTTTATATCAT 1960
 1621 GCAAAATATGTTTAAATAAATGAAATTTGTATT 1655
 1961 GCAAAATATGTTTAAATAAATGAAATTTGTATT 1995

RESULT 12
 AAV52576
 ID AAV52576 standard; DNA; 1997 BP.
 XX AC AAV52576;
 XX DT 14-DEC-1998 (first entry)
 XX DE Human vascular endothelial growth factor C gene.
 XX KW Flt4; vascular endothelial growth factor C; vascular endothelial cell;
 KW lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
 KW lymphangiogenesis; oedema; elephantiasis; Milroy's disease; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 352..1611
 XX FT /*tag= a
 XX FT /product= "VEGF-C protein"
 XX PN W09833917-A1.
 XX PD 06-AUG-1998.
 XX PF 02-FEB-1998; 98WO-US001973.
 XX PR 05-FEB-1997; 97US-00795430.

XX (LUDW-) LUDWIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX PI Alitalo K, Joukov V;
 XX WPI; 1998-437470/37.
 DR P-PSDB; AAW75740.
 XX New isolated vascular endothelial growth factor polypeptide(s) - used to
 PT develop products for treating, e.g. cancers, inflammation, oedema,
 PT granulocytopenia or for wound healing or tissue transplantation.
 XX Example 10; Page 112-115; 177pp; English.
 XX The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus
 XX Sequence 1997 BP; 526 A; 527 C; 490 G; 454 T; 0 U; 0 Other;

Query Match 98.7%; Score 1651.8; DB 2; Length 1997;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GTCCTTCCACCATGCTGCTGGCTTCTTCTGTGGGTGTCTCTGTCGCCGCTG 60
 DB 341 GTCCTTCCACCATGCTGCTGGCTTCTTCTGTGGGTGTCTCTGTCGCCGCTG 400
 QY 61 CGCTGCTCCGGGTCTCGGAGCGCCCGCCGCGCGCGCGCTTCGAGTCGGACTCG 120
 DB 401 CGCTGCTCCGGGTCTCGGAGCGCCCGCCGCGCGCGCTTCGAGTCGGACTCG 460
 QY 121 ACCTCTCGGAGCGCGAGCCCGAGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
 DB 461 ACCTCTCGGAGCGCGAGCCCGAGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 520
 QY 181 AGGAGCAGTTACGCTCTGTCTCAGTGTAGTGAATCATGACTGTACTCTACCCAGAT 240
 DB 521 AGGAGCAGTTACGCTCTGTCTCAGTGTAGTGAATCATGACTGTACTCTACCCAGAT 580
 QY 241 ATTGGAATATGTACAAGTGTACGTAAGGAAAGAGGCTGGCAACATAACAGAGAACAG 300
 DB 581 ATTGGAATATGTACAAGTGTACGTAAGGAAAGAGGCTGGCAACATAACAGAGAACAG 640
 QY 301 CCAACTCTAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGACATTAATATACAG 360
 DB 641 CCAACTCTAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGACATTAATATACAG 700
 QY 361 AGATCTTGAAGTATTTGATTAATGCTGGAGAGAGACTCAATGCTATGCCAGCGGAGTGT 420
 DB 701 AGATCTTGAAGTATTTGATTAATGCTGGAGAGAGACTCAATGCTATGCCAGCGGAGTGT 760
 QY 421 GTATAGATGTGGGAAAGAGTTTGGAGTGGCAACACACCTTCTTTTAAACCTCCATGTG 480
 DB 761 GTATAGATGTGGGAAAGAGTTTGGAGTGGCAACACACCTTCTTTTAAACCTCCATGTG 820
 QY 481 TGTCCGTCTACAGATGTGGGAGTGTGTCGAATAGTAGGGGTGCAAGTGAACACCA 540
 DB 821 TGTCCGTCTACAGATGTGGGAGTGTGTCGAATAGTAGGGGTGCAAGTGAACACCA 880
 QY 541 GCACGAGCTACCTCAGCAAGACGTTATTTTGAATTAACAGTCTCTCTCTCAAGGCCCA 600
 DB 881 GCACGAGCTACCTCAGCAAGACGTTATTTTGAATTAACAGTCTCTCTCTCTCAAGGCCCA 940

Qy	601	AACAGTAAGCAATCAGTTTTGCCAAATCACACTTCTCCGATGATGTCTAAACTGATG	660
Db	941	AACAGTAAGCAATCAGTTTTGCCAAATCACACTTCTCCGATGATGTCTAAACTGATG	1000
Qy	661	TTTACAGACAAGTTTCATTCCATTATTAGACGTTCCCTGCCAGCAAACTACCACAGTGC	720
Db	1001	TTTACAGACAAGTTTCATTCCATTATTAGACGTTCCCTGCCAGCAAACTACCACAGTGC	1060
Qy	721	AGGCAGCGAAACAAGACCTGCCCAACCAATATACATGTGGAATATCATCTGCAGATGCC	780
Db	1061	AGGCAGCGAAACAAGACCTGCCCAACCAATATACATGTGGAATATCATCTGCAGATGCC	1120
Qy	781	TGCGTCAAGGAAGATTTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGAATCC	840
Db	1121	TGCGTCAAGGAAGATTTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGAATCC	1180
Qy	841	ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACTGTGTCAGTGTCTGCGACAG	900
Db	1181	ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACTGTGTCAGTGTCTGCGACAG	1240
Qy	901	CGGGCTTCGGCTCGCCAGCTGTGGACCCCAACAAGAACTAGACAGAAAACTCATGCCAGT	960
Db	1241	CGGGCTTCGGCTCGCCAGCTGTGGACCCCAACAAGAACTAGACAGAAAACTCATGCCAGT	1300
Qy	961	GTGTCGTGTAACCAACAACTCTTCCCAGCCAAATGTGGGGCCAAACCGAGAATTTGATGAAA	1020
Db	1301	GTGTCGTGTAACCAACAACTCTTCCCAGCCAAATGTGGGGCCAAACCGAGAATTTGATGAAA	1360
Qy	1021	ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCCAGAAAAATCAACCCCTAAATCCTGGAA	1080
Db	1361	ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCCAGAAAAATCAACCCCTAAATCCTGGAA	1420
Qy	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTGTTAAAGGAAGAAGTTCC	1140
Db	1421	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTGTTAAAGGAAGAAGTTCC	1480
Qy	1141	ACCACCAAAATGCAGCTGTGTACAGACGGCCATGTACGAACCGCCAGAAAGCTGTGTAGC	1200
Db	1481	ACCACCAAAATGCAGCTGTGTACAGACGGCCATGTACGAACCGCCAGAAAGCTGTGTAGC	1540
Qy	1201	CAGGATTTTTCATATAGTGAAGAGTGTGCTGTGTGTCCTTCATATTTGGCAAGACCA	1260
Db	1541	CAGGATTTTTCATATAGTGAAGAGTGTGCTGTGTGTCCTTCATATTTGGAAAGACCA	1600
Qy	1261	AAATGAGCTAAGATTGTACTGTTTTTCAGTTTCATPCGATTTTCTATTATGAAAACTGTGT	1320
Db	1601	AAATGAGCTAAGATTGTACTGTTTTTCAGTTTCATPCGATTTTCTATTATGAAAACTGTGT	1660
Qy	1321	TGCCACAGTAGAACTGTCTGTGAACAGAGAGACCCCTTGTGGGTCCATGCTTAAACAAAGACA	1380
Db	1661	TGCCACAGTAGAACTGTCTGTGAACAGAGAGACCCCTTGTGGGTCCATGCTTAAACAAAGACA	1720
Qy	1381	AAAGTGTGCTTTTCCCTGAAACCATGTGGATTAACCTTTACAGAAATGGACTGGAGCTCATCTG	1440
Db	1721	AAAGTGTGCTTTTCCCTGAAACCATGTGGATTAACCTTTACAGAAATGGACTGGAGCTCATCTG	1780
Qy	1441	CAAAAGGCCCTCTGTGAAGACTGGTTTTCTGCGCAATGACCAAAACAGACCAAGATTTTCCCTC	1500
Db	1781	CAAAAGGCCCTCTGTGAAGACTGGTTTTCTGCGCAATGACCAAAACAGACCAAGATTTTCCCTC	1840
Qy	1501	TTGTGATTTCTTTAAAAAGATGACTATATAATTTATTTTCCACTAAAAATATGTTTCTGC	1560
Db	1841	TTGTGATTTCTTTAAAAAGATGACTATATAATTTATTTTCCACTAAAAATATGTTTCTGC	1900
Qy	1561	ATTTCATTTTATACCAACAACATTTGTTAAAACTCAGTGTGATCAATATTTTATATCAT	1620
Db	1901	ATTTCATTTTATACCAACAACATTTGTTAAAACTCAGTGTGATCAATATTTTATATCAT	1960
Qy	1621	GCAAAAATATGTTTAAAAATAAATGAAAAATGTATT	1655
Db	1961	GCAAAAATATGTTTAAAAATAAATGAAAAATGTATT	1995

XX	Sequence	1997 BP; 526 A; 527 C; 490 G; 454 T; 0 U; 0 Other;			
SY	Query Match	98.7%; Score 1651.8; DB 3; Length 1997;			
	Best Local Similarity	99.9%; Pred. No. 0;			
	Matches 1653; Conservative	0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	GTCTTCCACCATGCACTCGCTGGGTTCTCTCTGTGGGGTCTCTCTCTGGCGCTG 60	Db	1301	GTCTCTGTAAAAACAACACTTCTCCCGAGCCAATGTGGGGCCAAACGAGAAATTTGATGAA 1360
Db	341	GTCTTCCACCATGCACTCGCTGGGTTCTCTCTGTGGGGTCTCTCTCTGGCGCTG 400	QY	1021	ACATGCCAGTGTGTATGTAAAAAGAACCTCCCGCAGAAATCAACCCCTAAATCTCGAA 1080
QY	61	CGTGTCTCCGGTCTCTCGAGAGCGCCCGCCCGCCCGCTTCGAGTCCGACTCG 120	Db	1361	ACATGCCAGTGTGTATGTAAAAAGAACCTCCCGCAGAAATCAACCCCTAAATCTCGAA 1420
Db	401	CGTGTCTCCGGTCTCTCGAGAGCGCCCGCCCGCCCGCTTCGAGTCCGACTCG 460	QY	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGGAAAGTTC 1140
QY	121	ACCTCTCGAGCGGAGCGCGAGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 180	Db	1421	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGGAAAGTTC 1480
Db	461	ACCTCTCGAGCGGAGCGCGAGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 520	QY	1141	ACCACCAACATGCACTGTGTACAGAGCGGCGCATGTAGAAACCGCAGAGGCTTGTGAGC 1200
QY	181	AGGAGCAGTTACCGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAT 240	Db	1481	ACCACCAACATGCACTGTGTACAGAGCGGCGCATGTAGAAACCGCAGAGGCTTGTGAGC 1540
Db	521	AGGAGCAGTTACCGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAT 580	QY	1201	CAGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
QY	241	ATTGGAATGTACAGTGTACCTAAGGAAGAGGCTGGCAACATACAGAGAACAG 300	Db	1541	CAGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1600
Db	581	ATTGGAATGTACAGTGTACCTAAGGAAGAGGCTGGCAACATACAGAGAACAG 640	QY	1261	AAATGAGCTAAGATTTGTACTGTCTTCCAGTTTCATCGATTTTCTATTTATGGAATACTGTG 1320
QY	301	CCAACTCAACTCAAGCAGAGAGACTATAAATTTGCTGCAGCACATTAATAACAG 360	Db	1601	AAATGAGCTAAGATTTGTACTGTCTTCCAGTTTCATCGATTTTCTATTTATGGAATACTGTG 1660
Db	641	CCAACTCAACTCAAGCAGAGAGACTATAAATTTGCTGCAGCACATTAATAACAG 700	QY	1321	TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCCCTTGTGGTCCATGTCTAAACAGACA 1380
QY	361	AGATCTTGAAGATTTATGATGTGAGAAAGACTCAATGCATGCCACGGAGGTG 420	Db	1661	TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCCCTTGTGGTCCATGTCTAAACAGACA 1720
Db	701	AGATCTTGAAGATTTATGATGTGAGAAAGACTCAATGCATGCCACGGAGGTG 760	QY	1381	AAAGTCTGTCTTCTCCTGAACCATGTGGATAACTTTACAGAAATGGAGTGGAGCTCATCTG 1440
QY	421	GTATAGATGTGGGAGGAGTTTGGAGTCGCGACCAACACCTTCTTTAACTCCATGTG 480	Db	1721	AAAGTCTGTCTTCTCCTGAACCATGTGGATAACTTTACAGAAATGGAGTGGAGCTCATCTG 1780
Db	761	GTATAGATGTGGGAGGAGTTTGGAGTCGCGACCAACACCTTCTTTAACTCCATGTG 820	QY	1441	CAAAGGCTCTTGTGTAAAGACTGTGTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1500
QY	481	TGTCCGTCTACAGATGTGGGGTGTGTGCAATAGTGGAGGGCTGCAGTGCATGAACACA 540	Db	1781	CAAAGGCTCTTGTGTAAAGACTGTGTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1840
Db	821	TGTCCGTCTACAGATGTGGGGTGTGTGCAATAGTGGAGGGCTGCAGTGCATGAACACA 880	QY	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAATAATTTGTTCTGC 1560
QY	541	GCAGGAGCTACCTCAGCAGAGCTTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA 600	Db	1841	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAATAATTTGTTCTGC 1900
Db	881	GCAGGAGCTACCTCAGCAGAGCTTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA 940	QY	1561	ATTCATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATATCAT 1620
QY	601	AACAGTAACAATCAGTTTGTGCAATCAACACTTCTGCGGATGCATGTCTAACTGGATG 660	Db	1901	ATTCATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATATCAT 1960
Db	941	AACAGTAACAATCAGTTTGTGCAATCAACACTTCTGCGGATGCATGTCTAACTGGATG 1000			
QY	661	TTTACAGCAAGTTCAATCCATTTATAGACGTTCCCTGCGAGCAACACTACACAGTGTG 720			
Db	1001	TTTACAGCAAGTTCAATCCATTTATAGACGTTCCCTGCGAGCAACACTACACAGTGTG 1060			
QY	721	AGGAGCGAACAAGACTCTCCCAACCAATTCATGTGGAATATCAATCTGCAGATGCC 780			
Db	1061	AGGAGCGAACAAGACTCTCCCAACCAATTCATGTGGAATATCAATCTGCAGATGCC 1120			
QY	781	TGCTCAGGAAGATTTATGTTTCTCGATGTGGAGATGACTCAACAGATGGATTC 840			
Db	1121	TGCTCAGGAAGATTTATGTTTCTCGATGTGGAGATGACTCAACAGATGGATTC 1180			
QY	841	ATGACATCTGTGGACCAACAGAGGAGCTGGATGAAGAGACCTGTGAGTGTCTGCAGAG 900			
Db	1181	ATGACATCTGTGGACCAACAGAGGAGCTGGATGAAGAGACCTGTGAGTGTCTGCAGAG 1240			
QY	901	CGGGGCTTCGGCTGCGAGCTGTGGACCCCAACAAAGACTAGACAGAACTCATGCCAGT 960			
Db	1241	CGGGGCTTCGGCTGCGAGCTGTGGACCCCAACAAAGACTAGACAGAACTCATGCCAGT 1300			
QY	961	GTCTCTGTAAAAACAACACTTCTCCCGAGCCAATGTGGGGCCAAACCGAGAATTTGATGAA 1020			

RESULT 14

AAC62406
ID AAC62406 standard; cDNA; 1997 BP.

XX AAC62406;

XX 31-JAN-2001 (first entry)

XX Human VEGF-C coding sequence.

XX Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;
XX vascular endothelial growth factor receptor 3; VEGFR-3;
XX Milroy-Nonne syndrome; lymphoedema praecox; VEGF-C;
XX vascular endothelial growth factor C; ss.

OS Homo sapiens.

XX WO200058511-A1.

XX 05-OCT-2000.

XX 26-MAR-1999; 99WO-US006133.

XX 26-MAR-1999; 99WO-US006133.

XX (LUDW-) LUDWIG INST CANCER RES.

PA

XX	AAC68953;
XX	AC
XX	DT (first entry)
XX	Human VEGF-C coding sequence.
XX	
KW	Human; gene therapy; lymphatic disorder; hereditary lymphoedema; Flt4;
KW	vascular endothelial growth factor receptor-3; VEGFR-3; VEGF-C; VEGF-D;
XW	fms-like tyrosine kinase 4; ss.
OS	Homo sapiens.
XX	OS
XX	PN CA2283470-A1.
XX	PD 26-SEP-2000.
XX	PF 29-SEP-1999; 99CA-02283470.
XX	PR 26-MAR-1999; 99WG-US006133.
PR	16-AUG-1999; 99US-00375248.
XX	
PA	(UYPI-) UNIV PITTSBURGH.
PA	(UYHE-) UNIV HELSINKI LICENSING LTD OY.
PA	(LUDW-) LUDWIG INST CANCER RES.
PI	Alitalo K, Ferrell RE, Finsgold DN, Karkkainen M;
XX	
DR	WPI; 2001-007762/02.
DR	P-PSDB; AAB37605.
XX	
PT	Screening a human for an increased risk of developing lymphatic disorder
PT	comprises assaying nucleic acid for alterations in the sequences
PT	expressing vascular endothelial growth factor receptor-3.
XX	
PS	Disclosure; Page 59-61; 99pp; English.
XX	
CC	The present invention relates to a method for screening a human subject
CC	for an increased risk of developing a lymphatic disorder e.g. hereditary
CC	lymphoedema. The method comprises assaying nucleic acid of a human
CC	subject to determine a presence or an absence of a mutation altering the
CC	sequence or expression of vascular endothelial growth factor receptor-3
CC	(VEGFR-3)/fms-like tyrosine kinase 4 (Flt4) allele (see AAC68952 and
CC	AAB37604) and determining an increased risk of developing lymphatic
CC	disorder from presence or absence of the mutation. The presence of a
CC	mutation altering the encoded amino acid sequence or expression of at
CC	least 1 VEGFR-3 allele in the nucleic acid correlates with an increased
CC	risk of developing a lymphatic disorder. Treatment for hereditary
CC	lymphoedema can be provided through the administration of vascular
CC	endothelial growth factor C (VEGF-C) and vascular endothelial growth
CC	factor D VEGF-D genes (via gene therapy) and proteins. The present
CC	sequence is the coding sequence for VEGF-C
XX	
SQ	Sequence 1997 BP; 526 A; 527 C; 490 G; 454 T; 0 U; 0 Other;
	Query Match 98.7%; Score 1651.8; DB 4; Length 1997;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy	1 GTCCTTCCACCATGCAGTCGTGGCTTTCTCTGTGTGGCGTGTTCTGTCTGCCGCTG 60
Dd	341 GTCCCTTCACCATGCATGTCGTGGCTTTCTCTGTGTGGCGTGTTCTGTCTGCCGCTG 400
Oy	61 CGCTGCTCCCGGGTCTCTCGAGGCGCCCGCGCGCCCGCTTCGAGTCGGAATCG 120
Dd	401 CGTGCTCCCGGGTCTCTCGAGGCGCCCGCGCGCCCGCTTCGAGTCGGAATCG 460
Oy	121 ACCTCTCGAGCGCGAGCCCGACGCGGGCGAGGCCACCGCTTAGTCAGCAAGAATCGG 180
Dd	461 ACCTCTCGAGCGCGAGCCCGACGCGGGCGAGGCCACCGCTTAGTCAGCAAGAATCGG 520
Oy	181 AGGACAGATTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 240

Qy	1321	TGCCACAGTAGAACTGCTGTAACACAGAGAGACCCCTTGTGGGTCCATGCTAAACAAGACA	1380
Db	1661	TGCCACAGTAGAACTGCTGTAACACAGAGAGACCCCTTGTGGGTCCATGCTAAACAAGACA	1720
Qy	1381	AAAGTCTGCTTTTCCCTGAACCATGTGGATTAACCTTTACAGAAATGGAGCTGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTTTCCCTGAACCATGTGGATTAACCTTTACAGAAATGGAGCTGAGCTCATCTG	1780
Qy	1441	CAAAAGGCCCTTGTGTAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC	1500
Db	1781	CAAAAGGCCCTTGTGTAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC	1840
Qy	1501	TTGTGATTTCTTTAAAGAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC	1560
Db	1841	TTGTGATTTCTTTAAAGAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC	1900
Qy	1561	ATTCATTTTATAGCAACAACAAATGGTAAACTCACTGTGATCAATATTTTATATCAT	1620
Db	1901	ATTCATTTTATAGCAACAACAAATGGTAAACTCACTGTGATCAATATTTTATATCAT	1960
Qy	1621	GCAAAATATGTTTAAATAAAATGAAAAATGTATT	1655
Db	1961	GCAAAATATGTTTAAATAAAATGAAAAATGTATT	1995

Search completed: February 25, 2004, 19:10:37
Job time : 659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 16:56:30 ; Search time 147 Seconds
(without alignments)
5319.645 Million cell updates/sec

Title: US-09-921-143-1

Perfect score: 1674

Sequence: 1 gtcttcaccatgcactgc.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY NUC

Gapop 1070, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	100.0	1674	2	US-08-999-811-1
2	1674	100.0	1674	3	US-09-042-105-1
3	1674	100.0	1674	4	US-08-485-968-1
4	1664.6	99.4	2031	4	US-08-706-054A-1
5	1664.6	99.4	2031	4	US-08-313-299-1
6	1663	99.3	2031	4	US-08-706-054A-2
7	1663	99.3	2031	4	US-09-313-299-2
8	1661.2	99.2	1674	5	PCT-US96-09001-1
9	1651.8	98.7	1997	3	US-08-795-430-7
10	1651.8	98.7	1997	3	US-08-510-133A-34
11	1651.8	98.7	1997	4	US-09-355-700-7
12	1651.8	98.7	1997	4	US-08-601-132-32
13	1651.8	98.7	1997	4	US-08-671-573B-32
14	1526	91.2	1526	2	US-08-999-811-3
15	1526	91.2	1526	3	US-09-042-105-3
16	1520.2	90.8	1836	2	US-08-824-996-1
17	1192.2	70.6	1836	3	US-08-795-430-10
18	1182.2	70.6	1836	4	US-09-355-700-10
19	1182.2	70.6	1836	4	US-08-601-132-40
20	1182.2	70.6	1836	4	US-08-671-573B-40
21	1138.4	68.0	1140	3	US-08-510-133A-32
22	1138.4	68.0	1140	3	US-08-585-895-32
23	804.2	48.0	1741	3	US-08-795-430-12
24	804.2	48.0	1741	3	US-09-355-700-12
25	298	17.8	299	4	US-08-706-054A-6
26	298	17.8	299	4	US-09-313-299-6
27	144	8.6	1135	3	US-08-915-795-7

28	144	8.6	1325	3	US-08-915-795-6	Sequence 6, Appli
29	138.4	8.3	219	3	US-08-795-430-6	Sequence 6, Appli
30	138.4	8.3	219	3	US-08-510-133A-25	Sequence 25, Appli
31	138.4	8.3	219	3	US-08-585-895-25	Sequence 25, Appli
32	138.4	8.3	219	4	US-09-355-700-6	Sequence 6, Appli
33	138.4	8.3	219	4	US-08-601-132-25	Sequence 25, Appli
34	138.4	8.3	219	4	US-08-671-573B-25	Sequence 25, Appli
35	130.6	7.8	2846	3	US-08-915-795-1	Sequence 1, Appli
36	129	7.7	1337	3	US-09-016-434-982	Sequence 982, App
37	129	7.7	2029	3	US-08-915-795-4	Sequence 4, Appli
38	77	4.6	229	4	US-09-023-655-828	Sequence 828, App
39	51.8	3.1	7218	1	US-08-232-463-14	Sequence 14, Appli
40	50	3.0	50	4	US-08-706-054A-7	Sequence 7, Appli
41	50	3.0	50	4	US-08-706-054A-8	Sequence 8, Appli
42	50	3.0	50	4	US-09-313-299-7	Sequence 7, Appli
43	50	3.0	50	4	US-09-313-299-8	Sequence 8, Appli
44	48.4	2.9	495	4	US-09-244-583-25	Sequence 25, Appli
45	48.4	2.9	627	4	US-09-244-583-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1
US-08-999-811-1
; Sequence 1, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,811
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/207,550
; FILING DATE: 8-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWICZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1674 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 12..80
; FEATURE:

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/042,105
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/207,550
 ; FILING DATE: 8-MAR-1994
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/465,968
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: TO BE ASSIGNED
 ; FILING DATE: 24-DEC-1997
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ERIC K. STEFFE
 ; REGISTRATION NUMBER: 36,698
 ; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)371-2600
 ; TELEFAX: (202)371-2540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1674 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 12..80
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 81..1268
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 12..1268
 ; US-09-042-105-1

Query Match 100.0%; Score 1674; DB 3; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCCTCCACCATGCACTCGCTGGCTTCTCTCTGCGGTGTTCTCTGCTCGCGCTG 60
 Db 1 GTCCTCCACCATGCACTCGCTGGCTTCTCTCTGCGGTGTTCTCTGCTCGCGCTG 60
 QY 61 CGCTGCTCCGGGTCCTCGGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG 120
 Db 61 CGCTGCTCCGGGTCCTCGGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG 120
 QY 121 ACCTCTCGGAGCGGAGCCCGAGCGGCGAGCCCGCTTATGCAAGCAAGATCTGG 180
 Db 121 ACCTCTCGGAGCGGAGCCCGAGCGGCGAGCCCGCTTATGCAAGCAAGATCTGG 180
 QY 181 AGGAGCAGTACGCTGTGTCAGTGTAGATCACTGACTGACTCTACCCAGAT 240
 Db 181 AGGAGCAGTACGCTGTGTCAGTGTAGATCACTGACTGACTCTACCCAGAT 240
 QY 241 ATTGGAAATGTACAAGTGTCACTAAGGAAAGAGGCTGGCAACATACAGAACACAGG 300
 Db 241 ATTGGAAATGTACAAGTGTCACTAAGGAAAGAGGCTGGCAACATACAGAACACAGG 300
 QY 301 CCAACCTCACTCAGGACGAGAGACTATAAATTTGCTGCGACACATTATATACAG 360
 Db 301 CCAACCTCACTCAGGACGAGAGACTATAAATTTGCTGCGACACATTATATACAG 360
 QY 361 AGATCTTGAAAGATTGATAATGAGTGGAGAAAGACTCAATGATGCGGAGGTTGT 420
 Db 361 AGATCTTGAAAGATTGATAATGAGTGGAGAAAGACTCAATGATGCGGAGGTTGT 420

QY 421 GTATAGATGGGGAAGAGTTTGGAGTCGCGACAAACACCTTCTTAAACCTCCATGTG 480
 Db 421 GTATAGATGGGGAAGAGTTTGGAGTCGCGACAAACACCTTCTTAAACCTCCATGTG 480
 QY 481 TGTCCGCTTACAGATGTGGGGTTGTGCAATAGTAGAGGGGTGAGTGCATGAACACA 540
 Db 481 TGTCCGCTTACAGATGTGGGGTTGTGCAATAGTAGAGGGGTGAGTGCATGAACACA 540
 QY 541 GCACGAGTACCTCAGCAGAGAGTATTGAAATACAGTGCCTCTCTCAAGGCCCA 600
 Db 541 GCACGAGTACCTCAGCAGAGAGTATTGAAATACAGTGCCTCTCTCAAGGCCCA 600
 QY 601 AACAGTAACAATCAGTTTGGCAATCACTCTCCGCGATGATGTCTAAACTGATG 660
 Db 601 AACAGTAACAATCAGTTTGGCAATCACTCTCCGCGATGATGTCTAAACTGATG 660
 QY 661 TTTACAGACAAAGTTTATTCATTTAGACGTTTCCCTGCCAGCAACACTTACCAGTGT 720
 Db 661 TTTACAGACAAAGTTTATTCATTTAGACGTTTCCCTGCCAGCAACACTTACCAGTGT 720
 QY 721 AGGAGCGAAACAGACCTGCCCAACCAATTAATGTGGAAATACATCTCAGATGCC 780
 Db 721 AGGAGCGAAACAGACCTGCCCAACCAATTAATGTGGAAATACATCTCAGATGCC 780
 QY 781 TGGCTCAGGAAGATTATTTATGTTTCTCGATGCTGGAGTCACTCAACAGATGATCC 840
 Db 781 TGGCTCAGGAAGATTATTTATGTTTCTCGATGCTGGAGTCACTCAACAGATGATCC 840
 QY 841 ATGACATCTGTGACCAACAAAGAGCTGGATGAAGAGACCTGTGAGTGTCTCAGAG 900
 Db 841 ATGACATCTGTGACCAACAAAGAGCTGGATGAAGAGACCTGTGAGTGTCTCAGAG 900
 QY 901 CGGGCTTCGGCTGCGAGCTGTGAGCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
 Db 901 CGGGCTTCGGCTGCGAGCTGTGAGCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
 QY 961 GTGCTGTAAAAACAAACTCTTCCCGAGCCAATGTGGGCCCAACCGAGAAATTTGATGAAA 1020
 Db 961 GTGCTGTAAAAACAAACTCTTCCCGAGCCAATGTGGGCCCAACCGAGAAATTTGATGAAA 1020
 QY 1021 ACACATGCCAGTGTATGTAAAAAGAACTGCCCGCAAGAACTCAACCTTAAATCTCGAA 1080
 Db 1021 ACACATGCCAGTGTATGTAAAAAGAACTGCCCGCAAGAACTCAACCTTAAATCTCGAA 1080
 QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGATTC 1140
 Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGATTC 1140
 QY 1141 ACCACCAACATGCAGCTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
 Db 1141 ACCACCAACATGCAGCTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
 QY 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGTCCTTTCATATTGGGAAAGATCTGT 1260
 Db 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGTCCTTTCATATTGGGAAAGATCTGT 1260
 QY 1261 AATAGCTAAGATTTCTACTGTTTTCAGTTTTCATGATTTTCTATTGGGAAAGATCTGT 1320
 Db 1261 AATAGCTAAGATTTCTACTGTTTTCAGTTTTCATGATTTTCTATTGGGAAAGATCTGT 1320
 QY 1321 TGCCACAGTAGAAGTGTGTGTGAACAGAGACCCCTTGTGGGTCCATGCTAACAGAGACA 1380
 Db 1321 TGCCACAGTAGAAGTGTGTGTGAACAGAGACCCCTTGTGGGTCCATGCTAACAGAGACA 1380
 QY 1381 AAAGTGTCTTCTTCCCTGAACCATGTGATAATTTTACAGAAATGAGCTGGAGCTCATCTG 1440
 Db 1381 AAAGTGTCTTCTTCCCTGAACCATGTGATAATTTTACAGAAATGAGCTGGAGCTCATCTG 1440
 QY 1441 CAAAGGCTCTTGTAAAGACTGGTTTTCTGCAATGACCAACAGCCAGGATTTTCTC 1500
 Db 1441 CAAAGGCTCTTGTAAAGACTGGTTTTCTGCAATGACCAACAGCCAGGATTTTCTC 1500
 QY 1501 TTGTGATTTCTTTAAAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTGTTCTG 1560

```
Db 1501 TTGTGATTTCTTAAAGAAATGACTATATATATTTTCCACTAAAAATATTTGTTCTGC 1560
Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Qy 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTATATATATATATATATATATATATAT 1674
Db 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTATATATATATATATATATATATATAT 1674

RESULT 3
US-08-465-968-1
; Sequence 1, Application US/08465968E
; Patent No. 6608182
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor 2
; FILE REFERENCE: PF112PI
; CURRENT APPLICATION NUMBER: US/08/465,968E
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/207,550
; EARLIER FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1268)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (12)..(149)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (150)..(1268)
US-08-465-968-1

Query Match 100.0%; Score 1674; DB 4; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGGTTCCTCTGCTGCGCGCTG 60
Db 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGGTTCCTCTGCTGCGCGCTG 60

Qy 61 CGTGTCTCCGGGTCTCTCGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 120
Db 61 CGTGTCTCCGGGTCTCTCGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 120

Qy 121 ACCTCTCGAGCGGAGCGCGAGCGGGGCGAGCCAGCGCTTATCGAAGCAAGATCTGG 180
Db 121 ACCTCTCGAGCGGAGCGCGAGCGGGGCGAGCCAGCGCTTATCGAAGCAAGATCTGG 180

Qy 181 AGGAGCAGTTACGGTCTGTGTCAGGTAGATGAATCATGACTGTACTTACCCAGAAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCAGGTAGATGAATCATGACTGTACTTACCCAGAAT 240

Qy 241 ATTGGAATATGACAGTGTAGCTAAGGAAGAGGCTGGCAACATACAGAGACAGG 300
Db 241 ATTGGAATATGACAGTGTAGCTAAGGAAGAGGCTGGCAACATACAGAGACAGG 300

Qy 301 CCAACTCAACTCAAGCAGAGAGACTATAAAATTTGCTGCAGCACAATATATATACAG 360
Db 301 CCAACTCAACTCAAGCAGAGAGACTATAAAATTTGCTGCAGCACAATATATATACAG 360

Qy 361 AGATCTTGAAGATATGATATATGAGTGGAGAAAGACTCAATGCAATGCCAGGGAGTGT 420
Db 361 AGATCTTGAAGATATGATATATGAGTGGAGAAAGACTCAATGCAATGCCAGGGAGTGT 420
```

```
Qy 421 GTATAGATGTGGGAAGGAGTTTGGAGTGGCGACAAAACACCTTCTTTAAACCTCATGTG 480
Db 421 GTATAGATGTGGGAAGGAGTTTGGAGTGGCGACAAAACACCTTCTTTAAACCTCATGTG 480

Qy 481 TGTCCGTCTACAGATGTGGGGGTGTCTGCAATAGTAGGGGCTGCAGTGCATGAACCA 540
Db 481 TGTCCGTCTACAGATGTGGGGGTGTCTGCAATAGTAGGGGCTGCAGTGCATGAACCA 540

Qy 541 GCACGAGCTACCTCAGCAAGACGTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Db 541 GCACGAGCTACCTCAGCAAGACGTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600

Qy 601 AACCACTAACATCAGTTTGGCAATCAACATTTCTGCGGATGCATGTCTAAACTGGATG 660
Db 601 AACCACTAACATCAGTTTGGCAATCAACATTTCTGCGGATGCATGTCTAAACTGGATG 660

Qy 661 TTTACAGACAAGTTCAATTCATTTAGACGTTCCCTGCCAGCAACACTACCACAGTGC 720
Db 661 TTTACAGACAAGTTCAATTCATTTAGACGTTCCCTGCCAGCAACACTACCACAGTGC 720

Qy 721 AGGACGCAACAGAGCTGCCCAACCAATTCATGTGGAATATACATCTGCAGATGCC 780
Db 721 AGGACGCAACAGAGCTGCCCAACCAATTCATGTGGAATATACATCTGCAGATGCC 780

Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATCACTCAACAGATGGATCC 840
Db 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATCACTCAACAGATGGATCC 840

Qy 841 ATGACATCTGTGGACCAACCAAGAGCTGGATGAGAGAGCCTGTAGTGTCTGCAGAG 900
Db 841 ATGACATCTGTGGACCAACCAAGAGCTGGATGAGAGAGCCTGTAGTGTCTGCAGAG 900

Qy 901 CGGGGCTTCGGCTGCGAGCTGTGGAACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT 960
Db 901 CGGGGCTTCGGCTGCGAGCTGTGGAACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT 960

Qy 961 GTGTCTGTAAAAACAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
Db 961 GTGTCTGTAAAAACAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020

Qy 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTAAATCTCTGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTAAATCTCTGAA 1080

Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATCTGTTTAAAGAAAGAGTTC 1140
Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATCTGTTTAAAGAAAGAGTTC 1140

Qy 1141 ACCACCAACATCAGCTGTACAGAGCGCATGTACGAACCGCCAGAGGCTTTGTGAGC 1200
Db 1141 ACCACCAACATCAGCTGTACAGAGCGCATGTACGAACCGCCAGAGGCTTTGTGAGC 1200

Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCTCCCTTCATATTTGGCAAGACCAC 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCTCCCTTCATATTTGGCAAGACCAC 1260

Qy 1261 AATGAGCTAAGATTTGATCTGTTTCCAGTTCATCGATTTCTTATTTGAAACCTGTGT 1320
Db 1261 AATGAGCTAAGATTTGATCTGTTTCCAGTTCATCGATTTCTTATTTGAAACCTGTGT 1320

Qy 1321 TGCCACAGTGAAGCTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGTCTAAACAGACA 1380
Db 1321 TGCCACAGTGAAGCTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGTCTAAACAGACA 1380

Qy 1381 AAGTCTGTCTTCTGAAACCATGTGGATAAATTTACAGAAATGGAGTGCAGTCTCATCTG 1440
Db 1381 AAGTCTGTCTTCTGAAACCATGTGGATAAATTTACAGAAATGGAGTGCAGTCTCATCTG 1440

Qy 1441 CAAAGGCCCTCTGTGAAGAGTGTGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTC 1500
Db 1441 CAAAGGCCCTCTGTGAAGAGTGTGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTC 1500

Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATATTTTCCACTAAAAATATTTGTTCTGC 1560
```

Db 1501 TTGTGATTTCTTTAAAGAGATGACTATATATTTATTTTCCACTAAATAATTTCTTCTGC 1560
 Qy 1561 ATTCAATTTTATAGCAACAACAAATGGTAAAGACTCACTGTGATCAATATTTTATATCAT 1620
 Db 1561 ATTCAATTTTATAGCAACAACAAATGGTAAAGACTCACTGTGATCAATATTTTATATCAT 1620
 Qy 1561 GCAAAATATGTTTAAATATAAATGAAATTTGTATTTATAAATAAATAAATAA 1674
 Db 1561 GCAAAATATGTTTAAATATAAATGAAATTTGTATTTATAAATAAATAAATAA 1674

RESULT 4

US-08-706-054A-1
 ; Sequence 1, Application US/08705054A
 ; Patent No. 6451764
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, James
 ; Wood, William I.
 ; TITLE OF INVENTION: VEGF-Related Protein
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/706,054A
 ; FILING DATE: 30-Aug-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/003491
 ; FILING DATE: 08-Sep-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: P-40,378
 ; REFERENCE/DOCKET NUMBER: P0963R1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2031 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-08-706-054A-1

Query Match 99.4%; Score 1664.6; DB 4; Length 2031;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GTCTTCCACCATGACATCGTGGGCTTCTCTGTGGGCTGTCTCTGCTCGCGCTG 60
 Db 361 GTCTTCCACCATGACATCGTGGGCTTCTCTGTGGGCTGTCTCTGCTCGCGCTG 420
 Qy 61 CGCTGCTCCCGGTCCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 421 CGCTGCTCCCGGTCCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 Qy 121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 Db 481 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

Qy 181 AGGAGCAGTTACGCTCTGTGTCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAT 240
 Db 541 AGGAGCAGTTACGCTCTGTGTCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAT 600
 Qy 241 ATTGGAATAATGTACAAGTGTGAGCTTAAGGAAAGAGGCTGGCAACATACAGAGAAAGG 300
 Db 601 ATTGGAATAATGTACAAGTGTGAGCTTAAGGAAAGAGGCTGGCAACATACAGAGAAAGG 660
 Qy 301 CCAACCTCAACTCAAGGACAGAGACTATAAATTTGCTGTGAGCAGATTAATATACAG 360
 Db 661 CCAACCTCAACTCAAGGACAGAGACTATAAATTTGCTGTGAGCAGATTAATATACAG 720
 Qy 361 AGATCTTGAAGAATGATTAATGATGAGTGAAGAGACTCAATGATGATGCAACGAGGCTG 420
 Db 721 AGATCTTGAAGAATGATTAATGATGAGTGAAGAGACTCAATGATGATGCAACGAGGCTG 780
 Qy 421 GTATAGATGTGGGAAAGAGTTTGGAGTCGCAACAAACACCTTCTTTAAACCTCATGTG 480
 Db 781 GTATAGATGTGGGAAAGAGTTTGGAGTCGCAACAAACACCTTCTTTAAACCTCATGTG 840
 Qy 481 TGTCCGCTCTACAGATGTGGGCTTGTCTCAATAGTGAAGGCTGCAGTGCATGAACACCA 540
 Db 841 TGTCCGCTCTACAGATGTGGGCTTGTCTCAATAGTGAAGGCTGCAGTGCATGAACACCA 900
 Qy 541 GCACGAGCTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
 Db 901 GCACGAGCTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 960
 Qy 601 AACAGTAACAAATGATTTTGCCTCAATGATGAGGCTGCAGTGCATGAACACCA 660
 Db 961 AACAGTAACAAATGATTTTGCCTCAATGATGAGGCTGCAGTGCATGAACACCA 1020
 Qy 661 TTTACAGACAAGTTTCATTTCCATTTATTAGACGTTCCCTGCCAGCAACACTACACAGTGC 720
 Db 1021 TTTACAGACAAGTTTCATTTCCATTTATTAGACGTTCCCTGCCAGCAACACTACACAGTGC 1080
 Qy 721 AGGAGCGAAACAAGACCTGCCCCACCAATTAATGATGGAATTAATCAATCTGAGATGCC 780
 Db 1081 AGGAGCGAAACAAGACCTGCCCCACCAATTAATGATGGAATTAATCAATCTGAGATGCC 1140
 Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
 Db 1141 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 1200
 Qy 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTCAAGTGTCTGCAGAG 900
 Db 1201 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTCAAGTGTCTGCAGAG 1260
 Qy 901 CGGGCTTCGGCTCGCAGCTGTGGACCCCAAGAGACTAGACAGAACTCATGCCAGT 960
 Db 1261 CGGGCTTCGGCTCGCAGCTGTGGACCCCAAGAGACTAGACAGAACTCATGCCAGT 1320
 Qy 961 GTGCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGAAATTTGATGAA 1020
 Db 1321 GTGCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGAAATTTGATGAA 1380
 Qy 1021 ACATGCGCAGTGTGTATGTAAGAACTCCCGCAGAAATCAACCCCTTAATCTCGAA 1080
 Db 1381 ACATGCGCAGTGTGTATGTAAGAACTCCCGCAGAAATCAACCCCTTAATCTCGAA 1440
 Qy 1081 AATGTGCTGTGAATGTACAGAACTCCACAGAAATGCTTGTAAAAAGAAAGATTC 1140
 Db 1441 AATGTGCTGTGAATGTACAGAACTCCACAGAAATGCTTGTAAAAAGAAAGATTC 1500
 Qy 1141 ACCACCAACATGACGCTGTTACAGACGCGCATGTAACGCGCAGAGGCTTGTGAGC 1200
 Db 1501 ACCACCAACATGACGCTGTTACAGACGCGCATGTAACGCGCAGAGGCTTGTGAGC 1560
 Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTCCCTTCATTTGGGAAAGACAC 1260
 Db 1561 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTCCCTTCATTTGGGAAAGACAC 1620
 Qy 1261 AAATGAGCTAAGATTTGCTGTTTTCCAGTTTCATGATTTTTCTATATATGGAATGCTGT 1320

Db 1621 AAATGAGCTAAGATTGTACTGTGTTTCCAGTTCATCGATTTCTATTATGGAAAACTGTGT 1680
Qy 1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCCCTTGTGGGTCCATGCTAACAAAGACA 1380
Db 1681 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCCCTTGTGGGTCCATGCTAACAAAGACA 1740
Qy 1381 AAAGTGTCTTCTCTGAACCAATGGGATACCTTTACAGAAATGGAGCTGAGCTCATCTG 1440
Db 1741 AAAGTGTCTTCTCTGAACCAATGGGATACCTTTACAGAAATGGAGCTGAGCTCATCTG 1800
Qy 1441 CAABAGGCTCTGTGAAGAGCTGTTTCTGCAATGACCAAAACAGCCCAAGATTTCTCTC 1500
Db 1801 CAABAGGCTCTGTGAAGAGCTGTTTCTGCAATGACCAAAACAGCCCAAGATTTCTCTC 1860
Qy 1501 TTGTGATTTCTTTAAAGAAATGACATATATAATTTATTTCCACTAAAAATATGTTTCTGC 1560
Db 1861 TTGTGATTTCTTTAAAGAAATGACATATATAATTTATTTCCACTAAAAATATGTTTCTGC 1920
Qy 1561 ATTCATTTTATAGCAACACAAATGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1921 ATTCATTTTATAGCAACACAAATGGTAAACTCACTGTGATCAATATTTTATATCAT 1980
Qy 1621 GCAAAATATGTTTAAATATAAATGAAAATGTAATTTATAAAAAA 1671
Db 1981 GCAAAATATGTTTAAATATAAATGAAAATGTAATTTATAAAAAA 2031

RESULT 5
US-09-313-299-1
; Sequence 1, Application US/09313299B
; Patent No. 6576608
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: VEGF-RELATED PROTEIN
; FILE REFERENCE: P0963R1D1
; CURRENT APPLICATION NUMBER: US/09/313,299B
; EARLIER FILING DATE: 1999-05-17
; EARLIER APPLICATION NUMBER: US 08/706,054
; EARLIER FILING DATE: 1996-08-30
; EARLIER APPLICATION NUMBER: US 60/003,491
; EARLIER FILING DATE: 1995-09-08
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 1
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Human
; LOCATION: 1-2031
; OTHER INFORMATION: Sequence source: VRP
; Patent No. 6576608
US-09-313-299-1

Query Match 99.4%; Score 1664.6; DB 4; Length 2031;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGCTTCCACATGCACTCGCTGGGCTTCTCTCTGTGGGCTTCTCTGTGGGCTTCTCTGTGGGCTG 60
Db 361 GTGCTTCCACATGCACTTCTGGGCTTCTCTCTGTGGGCTTCTCTGTGGGCTTCTCTGTGGGCTG 420
Qy 61 CGTGTCTCCGGTCTCTCGAGGCGCCCGCGCGCTTCGAGTCCGGAATCG 120
Db 421 CGTGTCTCCGGTCTCTCGAGGCGCCCGCGCGCTTCGAGTCCGGAATCG 480
Qy 121 ACCCTCGGACGCGGAGCCCGCGCGCGGCGAGGCGCGGCTTATGCAAGCAAGATCTGG 180
Db 481 ACCCTCGGACGCGGAGCCCGCGCGCGGCGAGGCGCGGCTTATGCAAGCAAGATCTGG 540
Qy 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGTGAATCATGACTGTACTCTACCCAGAAAT 240

Db 541 AGGAGCAGTTACGGTCTGTGTCAGTGTAGTGAATCATGACTGTACTCTACCCAGAAT 600
Qy 241 ATTGGAAAATGTACAAGTGTCAAGTGTAAAGAAAGAGGCTGGCAACATACACAGAACAGG 300
Db 601 ATTGGAAAATGTACAAGTGTCAAGTGTAAAGAAAGAGGCTGGCAACATACACAGAACAGG 660
Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATATAATTTGCTGCAGCACATTATAATACAG 360
Db 661 CCAACCTCAACTCAAGGACAGAGAGACTATATAATTTGCTGCAGCACATTATAATACAG 720
Qy 361 AGATCTTTGAAAAGTATTGATTAATGAGTGGAGAAAGCTCAATGCAATGCGCCAGGAGTGT 420
Db 721 AGATCTTTGAAAAGTATTGATTAATGAGTGGAGAAAGCTCAATGCAATGCGCCAGGAGTGT 780
Qy 421 GTATAGATGTGGGAAAGAGTGTGGAGTCCGACAAACACCTTCTTTAAACCTCCATGTG 480
Db 781 GTATAGATGTGGGAAAGAGTGTGGAGTCCGACAAACACCTTCTTTAAACCTCCATGTG 840
Qy 481 TGTCCGCTTACAGATGTGGGGTGTGTCATATAGTAGAGGGCTGCAGTGCATGAACACCA 540
Db 841 TGTCCGCTTACAGATGTGGGGTGTGTCATATAGTAGAGGGCTGCAGTGCATGAACACCA 900
Qy 541 GCAGGAGCTACCTCAGCAAGAGCTTATTTGAAATACAGTGCCTCTCTCAAGGCCCA 600
Db 901 GCAGGAGCTACCTCAGCAAGAGCTTATTTGAAATACAGTGCCTCTCTCAAGGCCCA 960
Qy 601 AACCAGTAAACAATCAGTTTTGCGCAATCACTTCTCGCGGATGATGTCTAACTGGATG 660
Db 961 AACCAGTAAACAATCAGTTTTGCGCAATCACTTCTCGCGGATGATGTCTAACTGGATG 1020
Qy 661 TTTACAGACAGAGTTCATTTCCATTTATAGAGTTCCTGCGAGCACTACACAGTGTG 720
Db 1021 TTTACAGACAGAGTTCATTTCCATTTATAGAGTTCCTGCGAGCACTACACAGTGTG 1080
Qy 721 AGGAGCGAAACAAGAGCTGCCCAACCAATTACATGTGGAATAATCACTGTGAGATGCC 780
Db 1081 AGGAGCGAAACAAGAGCTGCCCAACCAATTACATGTGGAATAATCACTGTGAGATGCC 1140
Qy 781 TGGCTCAGGAAGATTTATGTTTCTCGATGCTCGAGATGACTCAACAGATGGATTC 840
Db 1141 TGGCTCAGGAAGATTTATGTTTCTCGATGCTCGAGATGACTCAACAGATGGATTC 1200
Qy 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTCTCAGTGTGTGCGAG 900
Db 1201 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTCTCAGTGTGTGCGAG 1260
Qy 901 CGGGCTTCGGCTCGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1261 CGGGCTTCGGCTCGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 1320
Qy 961 GTGTCTGTAATAAACAACAACTCTTCCCGAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
Db 1321 GTGTCTGTAATAAACAACAACTCTTCCCGAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1380
Qy 1021 ACACATGCCAGTGTGATGTAAGAAGACCTGCCCGAGAAATCAACCCCTAAATCTCTGGAA 1080
Db 1381 ACACATGCCAGTGTGATGTAAGAAGACCTGCCCGAGAAATCAACCCCTAAATCTCTGGAA 1440
Qy 1081 AATGTGCTGTGATGATGACAGAAAGTCCACAGAAATGCTTTGTTAAAGAGGAAGATTC 1140
Db 1441 AATGTGCTGTGATGATGACAGAAAGTCCACAGAAATGCTTTGTTAAAGAGGAAGATTC 1500
Qy 1141 ACCACCAACATCAGCTGTGTACAGCGGCATGTACGAACCGCCAGAGGCTTGTGAGC 1200
Db 1501 ACCACCAACATCAGCTGTGTACAGCGGCATGTACGAACCGCCAGAGGCTTGTGAGC 1560
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGCTTCCATATATGCAAGAACAC 1260
Db 1561 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGCTTCCATATATGCAAGAACAC 1620
Qy 1261 AAATGAGCTAAGATTTGACTGTTTTTCCAGTTTCATGATTTTCTATATGGAATACTGTGT 1320
Db 1621 AAATGAGCTAAGATTTGACTGTTTTTCCAGTTTCATGATTTTCTATATGGAATACTGTGT 1680

1321 TGCCACAGTAGAAGTCTGTGTGAAACAGAGAGACCCCTGTGGGTCCATCTAACAAAGACA 1380
 1681 TGCCACAGTAGAAGTCTGTGTGAAACAGAGAGACCCCTGTGGGTCCATCTAACAAAGACA 1740
 1381 AAAGTCTGTCTTCTTCAACCAATGTGGATACTTTACAGAAATGAGCTGGAGCTCATCTG 1440
 1741 AAAGTCTGTCTTCTTCAACCAATGTGGATACTTTACAGAAATGAGCTGGAGCTCATCTG 1800
 1441 CAAAAGCCCTCTGTGTAAGAGCTGGTTTCTGCAATGACCCAAACAGCCAAAGATTTCTTC 1500
 1801 CAAAAGCCCTCTGTGTAAGAGCTGGTTTCTGCAATGACCCAAACAGCCAAAGATTTCTTC 1860
 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTTCCACTAAAAATATTGTTCTGTC 1560
 1861 TTGTGATTTCTTTAAAGAAATGACTATATAATTTTCCACTAAAAATATTGTTCTGTC 1920
 1561 ATTCAATTTTATAGCAACAACTTGTGTAAGAGCTGGTTTCTGCAATGACCCAAACAGCCAAAGATTTCTTC 1620
 1921 ATTCAATTTTATAGCAACAACTTGTGTAAGAGCTGGTTTCTGCAATGACCCAAACAGCCAAAGATTTCTTC 1980
 1621 GCAAAATATGTTTAAATATAATGAAATTTGTAATTTATAAAAAA 1671
 1981 GCAAAATATGTTTAAATATAATGAAATTTGTAATTTATAAAAAA 2031

RESULT 6

US-08-706-054A-2/c
 ; Sequence 2, Application US/08706054A
 ; Patent No. 6451764
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, James
 ; TITLE OF INVENTION: VEGF-Related Protein
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/706, 054A
 ; FILING DATE: 30-Aug-1996
 ; CLASSIFICATION: <unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/003491
 ; FILING DATE: 08-Sep-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: P-40, 378
 ; REFERENCE/DOCKET NUMBER: P0963R1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 ;
 ; Query Match 99.3%; Score 1663; DB 4; Length 2031;
 ; Best Local Similarity 99.7%; Pred. No. 0;
 ; US-08-706-054A-2

Matches 1666; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGGCTGTCTCTGTCCCGCTG 60
 DB 1671 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGGCTGTCTCTGTCCCGCTG 1612
 QY 61 CGCTCTCCCGGCTCCTCGGAGGCGCCGCGCGCGCGCGCGCGCTCGAGTCCGAGCTCG 120
 DB 1611 TGCTCTCCCGGCTCCTCGGAGGCGCCGCGCGCGCGCGCGCTCGAGTCCGAGCTCG 1552
 QY 121 ACCTTCGAGCGCGGAGCCCGACGCGGCGGAGGCTGCAAGCAAGATCTCG 180
 DB 1551 ACCTTCGAGCGCGGAGCCCGACGCGGCGGAGGCTGCAAGCAAGATCTCG 1492
 QY 181 AGGACGAGTTAGGCTCTGTGTCAGTGTAGATGAATCTCATGACTGTACTCTACCCAGAT 240
 DB 1491 AGGACGAGTTAGGCTCTGTGTCAGTGTAGATGAATCTCATGACTGTACTCTACCCAGAT 1432
 QY 241 ATTGAAAATGTACAAAGTGTGCTTAAGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
 DB 1431 ATTGAAAATGTACAAAGTGTGCTTAAGAAAGAGGCTGGCAACATAACAGAGAACAGG 1372
 QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGCATTTATAATACAG 360
 DB 1371 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGCATTTATAATACAG 1312
 QY 361 AGATCTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGATGCCACGAGGAGTGT 420
 DB 1311 AGATCTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGATGCCACGAGGAGTGT 1252
 QY 421 GTATAGATGTGGGAAAGGAGTTTGGAGTCGCGACAAACACCTTCTTAAACCTCCATGTG 480
 DB 1251 GTATAGATGTGGGAAAGGAGTTTGGAGTCGCGACAAACACCTTCTTAAACCTCCATGTG 1192
 QY 481 TGTCCGTCTACAGATGTGGGAGTTTGTCAATAGTGGAGGCTGCAATGATGAACACCA 540
 DB 1191 TGTCCGTCTACAGATGTGGGAGTTTGTCAATAGTGGAGGCTGCAATGATGAACACCA 1132
 QY 541 GCACGAGTCTACTCAGCAAGAGCTTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
 DB 1131 GCACGAGTCTACTCAGCAAGAGCTTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 1072
 QY 601 AACAGTAACAAATCAGTTTGGCAATCAGCTTCCCTCCGATGATGTCTAAATGATG 660
 DB 1071 AACAGTAACAAATCAGTTTGGCAATCAGCTTCCCTCCGATGATGTCTAAATGATG 1012
 QY 661 TTACAGACAAAGTTCATTTCCTATTTAGAGCTTCCCTCCGAGAGTACTCAACAGATGGATTC 720
 DB 1011 TTACAGACAAAGTTCATTTCCTATTTAGAGCTTCCCTCCGAGAGTACTCAACAGATGGATTC 952
 QY 721 AGGACGGAACAAAGACCTTCCCAATTAATGATGTAATTAATCAATCTGAGATGCC 780
 DB 951 AGGACGGAACAAAGACCTTCCCAATTAATGATGTAATTAATCAATCTGAGATGCC 892
 QY 781 TGGCTCAGAAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTC 840
 DB 891 TGGCTCAGAAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTC 832
 QY 841 ATGACATCTGTGAGCAACAAAGAGCTGATGAAGAGCTGTGAGTGTGTCTGAGAG 900
 DB 831 ATGACATCTGTGAGCAACAAAGAGCTGATGAAGAGCTGTGAGTGTGTCTGAGAG 772
 QY 901 CGGGCTTCCGCTCCGAGCTGTGAGCCCAAGAAAGTCTAGACAGAAATCTCATGCGAGT 960
 DB 771 CGGGCTTCCGCTCCGAGCTGTGAGCCCAAGAAAGTCTAGACAGAAATCTCATGCGAGT 712
 QY 961 GTGTCTGTAAAAACAAACTCTTCCCGAGCCATGTGGGGCCACCGAGATTTGATGAAA 1020
 DB 711 GTGTCTGTAAAAACAAACTCTTCCCGAGCCATGTGGGGCCACCGAGATTTGATGAAA 652
 QY 1021 ACACATGCCAGTGTGTATGTAAGAAAGACCTTCCCGAGCCATGTGGGGCCACCGAGATTTGATGAAA 1080
 DB 651 ACACATGCCAGTGTGTATGTAAGAAAGACCTTCCCGAGCCATGTGGGGCCACCGAGATTTGATGAAA 592


```

841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGAGTGTGTCTGCAGAG 900
901 CGGGGCTTGGGCTGGCTGGAGCTGGAGCCCAAGAGAACTAGACAGAACTCATGCCAGT 960
901 CGGGGCTTGGGCTGGCTGGAGCTGGAGCCCAAGAGAACTAGACAGAACTCATGCCAGT 960
961 GTGTCTGTAAAAACAACTCTTCCCGAGCAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
961 GTGTCTGTAAAAACAACTCTTCCCGAGCAATGTGGGCGCAACCGAGCAATTTGATGAAA 1020
1021 ACACATGCCAGTGTGTATGAAGAGAACTGGCCCGAGAAATCAACCCCTAAATCTGAAA 1080
1021 ACACATGCCAGTGTGTATGAAGAGAACTGGCCCGAGAAATCAACCCCTAAATCTGAAA 1080
1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAAATTC 1140
1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAAATTC 1140
1141 ACCACCAACATGCAGCTGTACAGAGCGGCATGTACGACCGCCAGAGGCTTGTGAGC 1200
1141 ACCACCAACATGCAGCTGTACAGAGCGGCATGTACGACCGCCAGAGGCTTGTGAGC 1200
1201 CAGGATTTTCAATAGTGAAGAGTGTGTGCTGCTCCCTTCAATATGCGCAAGACAC 1260
1201 CAGGATTTTCAATAGTGAAGAGTGTGTGCTGCTCCCTTCAATATGCGCAAGACAC 1260
1261 AATAGAGCTAAGATGTATGCTGTTTCCAGTTCATCGATTTTCTATATGGAAGAACTGT 1320
1261 AATAGAGCTAAGATGTATGCTGTTTCCAGTTCATCGATTTTCTATATGGAAGAACTGT 1320
1321 TGCCACAGTAGAAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1321 TGCCACAGTAGAAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1381 AATAGCTGTCTTCTGAACCATGTGATGATTTTACAGAAATGGAAGTGGAGCTCATCTG 1440
1381 AATAGCTGTCTTCTGAACCATGTGATGATTTTACAGAAATGGAAGTGGAGCTCATCTG 1440
1441 CAAAGGCTTCTTGAAGAGTGTGTGCTGCTCCCTTCAATATGCGCAAGAGAGAGAG 1500
1441 CAAAGGCTTCTTGAAGAGTGTGTGCTGCTCCCTTCAATATGCGCAAGAGAGAGAG 1500
1501 TTGTGATTTCTTAAAGAGTGTATATAATTTTCCACTTAAAGAAATTTGTTTCTG 1560
1501 TTGTGATTTCTTAAAGAGTGTATATAATTTTCCACTTAAAGAAATTTGTTTCTG 1560
1561 ATTCAATTTTATAGCAACAAATGTTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGT 1620
1561 ATTCAATTTTATAGCAACAAATGTTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGT 1620
1621 GCAAAATATCTTTAAAGTAAATGAAATGTAATTTTATATAAAGGAGAGAGAGAGAG 1674
1621 GCAAAATATCTTTAAAGTAAATGAAATGTAATTTTATATAAAGGAGAGAGAGAGAG 1674

```

```

RESULT 9
US-08-795-430-7
; Sequence 7, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/Fin96/00427
; FILING DATE: 01-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33691
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1997 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1608
; US-08-795-430-7

```

```

Query Match 98.7%; Score 1651.8; DB 3; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGGCTGTCTCTCTCGCCGCTG 60
DB 341 GTCCTTCCACCATGCACTTGTGGGCTTCTTCTGTGGGCTGTCTCTCTCGCCGCTG 400
QY 61 CGCTGCTCCCGGCTCCTCGGAGGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
DB 401 CGCTGCTCCCGGCTCCTCGGAGGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 460
QY 121 ACCTCTCGGAGCGCGAGCGCGAGCGCGAGCGCGAGCGCGCTTATGCAAGCAAGATCTGG 180
DB 461 ACCTCTCGGAGCGCGAGCGCGAGCGCGAGCGCGAGCGCGCTTATGCAAGCAAGATCTGG 520
QY 181 AGGAGCGATGAGGCTGTGTGTCCAGTGTAGATGAGTCACTGAGTGTACTTACCCAGAT 240
DB 521 AGGAGCGATGAGGCTGTGTGTCCAGTGTAGATGAGTCACTGAGTGTACTTACCCAGAT 580
QY 241 ATTGGAATAATGTACAAAGTGTGAGTAAAGAGAGAGCTGCAACATAACAGAGAAACAGG 300
DB 581 ATTGGAATAATGTACAAAGTGTGAGTAAAGAGAGAGCTGCAACATAACAGAGAAACAGG 640
QY 301 CCAACCTCAACTCAAGGACAGAGAGAGATATAAATTTGCTGAGGACATTTATATACAG 360

```

Db 641 CCAACCTCAACTCAAGACAGAGAGACTATAAAATTTGCTGCGACACATTAATAACAG 700
 Qy 361 AGATCTTTGAAAGATTGATTAATGAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 420
 Db 701 AGATCTTTGAAAGATTGATTAATGAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 760
 Qy 421 GTATAGATGTGGGAAAGAGTTTGGAGTGGGCAAAACACCTTCTTTAAACCTCCATGTG 480
 Db 761 GTATAGATGTGGGAAAGAGTTTGGAGTGGGCAAAACACCTTCTTTAAACCTCCATGTG 820
 Qy 481 TGTCCGCTCTACAGATGTGGGGTTGCTGCAATAGTAGGGGCTGCAATGATGAACACCA 540
 Db 821 TGTCCGCTCTACAGATGTGGGGTTGCTGCAATAGTAGGGGCTGCAATGATGAACACCA 880
 Qy 541 GCACGAGTACCTCAGCAAGAGCTTTAATGAAATTAAGTGCCTCTCTCAAGGCCCCA 600
 Db 881 GCACGAGTACCTCAGCAAGAGCTTTAATGAAATTAAGTGCCTCTCTCAAGGCCCCA 940
 Qy 601 AACCAGTAACCAATCAGTTTTCCTCCCAATCACAATCTCCGATGATGCTTAAACTGATG 660
 Db 941 AACCAGTAACCAATCAGTTTTCCTCCCAATCACAATCTCCGATGATGCTTAAACTGATG 1000
 Qy 661 TTTACAGCAAGTTCAATCCATTAATAGAGTTCCTCCGAGCAACACTACCAAGTGC 720
 Db 1001 TTTACAGCAAGTTCAATCCATTAATAGAGTTCCTCCGAGCAACACTACCAAGTGC 1060
 Qy 721 AGGAGGCAACAGACTGCCCAACCAATACATGTCGAATATACATCTGCAGATGCC 780
 Db 1061 AGGAGGCAACAGACTGCCCAACCAATACATGTCGAATATACATCTGCAGATGCC 1120
 Qy 781 TGGCTCAGAAAGATTTATGTTTTCCTCGATGCTGGAGATGACTCAACAGATGGATTCC 840
 Db 1121 TGGCTCAGAAAGATTTATGTTTTCCTCGATGCTGGAGATGACTCAACAGATGGATTCC 1180
 Qy 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTCACTGTCTGCGAG 900
 Db 1181 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTCACTGTCTGCGAG 1240
 Qy 901 CGGGCTTCGGCTCGCCAGCTGTGAGCCCAAGAACTAGACAGAACTCATGCCAGT 960
 Db 1241 CGGGCTTCGGCTCGCCAGCTGTGAGCCCAAGAACTAGACAGAACTCATGCCAGT 1300
 Qy 961 GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCCAACCGAGAAATTTGATGAAA 1020
 Db 1301 GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCCAACCGAGAAATTTGATGAAA 1360
 Qy 1021 ACACATGCGAGTGTATGTAAGAACTGCCCCAGAAATCAACCCCTAACTCCTGGA 1080
 Db 1361 ACACATGCGAGTGTATGTAAGAACTGCCCCAGAAATCAACCCCTAACTCCTGGA 1420
 Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
 Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1480
 Qy 1141 ACCACCAACATGAGCTGTATACAGCGGCAATGTACGACCGGCAAGGCTTGTGAGC 1200
 Db 1481 ACCACCAACATGAGCTGTATACAGCGGCAATGTACGACCGGCAAGGCTTGTGAGC 1540
 Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTCTGTTGTCCTTTCATATTTGGCAAGACAC 1260
 Db 1541 CAGGATTTTCATATAGTGAAGAGTGTCTGTTGTCCTTTCATATTTGGAAAGACAC 1600
 Qy 1261 AAATGAGCTAAGATTGTACTGTTTTCAGTTTCATGATTTTCTATTTGGAATCTGTGT 1320
 Db 1601 AAATGAGCTAAGATTGTACTGTTTTCAGTTTCATGATTTTCTATTTGGAATCTGTGT 1660
 Qy 1321 TGCCACAGTGAAGTGTCTGTGAACAGAGAGACCTTGTGGGTCCATGCTAACAAAGACA 1380
 Db 1661 TGCCACAGTGAAGTGTCTGTGAACAGAGAGACCTTGTGGGTCCATGCTAACAAAGACA 1720
 Qy 1381 AAAGTCTGTCTTCTGTAACCACTGTGGATAACTTTTACAGAAATGAGCTGGAGCTCATCTG 1440
 Db 1721 AAAGTCTGTCTTCTGTAACCACTGTGGATAACTTTTACAGAAATGAGCTGGAGCTCATCTG 1780

Qy 1441 CAAAAGGCTCTTGTAAAGACTGTTTTCTGCCAATGACCAAAACAGCCAGATTTTCTC 1500
 Db 1781 CAAAAGGCTCTTGTAAAGACTGTTTTCTGCCAATGACCAAAACAGCCAGATTTTCTC 1840
 Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1560
 Db 1841 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1900
 Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCCTGTGATCAATATTTTATATCAT 1620
 Db 1901 ATTCAATTTTATAGCAACAAATTTGGTAAACTCCTGTGATCAATATTTTATATCAT 1960
 Qy 1621 GCAAAATGTTTAAATAAATGAAATTTGTTT 1655
 Db 1961 GCAAAATGTTTAAATAAATGAAATTTGTTT 1995

RESULT 10

US-08-510-133A-34
 ; Sequence 34, Application US/08510133A
 ; Patent No. 6221839
 ; GENERAL INFORMATION:
 ; APPLICANT: Alitalo, Kari
 ; Joukov, Vladimir
 ; TITLE OF INVENTION: Receptor Ligand
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/510,133A
 ; FILING DATE: 01-Aug-1995
 ; CLASSIFICATION: <unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gass, David A.
 ; REGISTRATION NUMBER: 38,153
 ; REFERENCE/DOCKET NUMBER: 28113/32863
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0446
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1997 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 352..1608
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
 US-08-510-133A-34

Query Match 98.7%; Score 1651.8; DB 3; Length 1997;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GTCTTCCACCATGCACTCGTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 60
 Db 341 GTCTTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 400
 Qy 61 CGCTGTCTCCGGGTCTCTCGGAGGCGCCGCGCGCTTCGAGTCGGGACTCG 120

Db 401 CGCTGCTCCGGGTCCTCGAGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG 460
Qy 121 ACCTCTCGGACGGGAGCCCGCGCGCGGAGCCACGGCTTATGCAAGCAAGATCTCG 180
Db 461 ACCTCTCGGACGGGAGCCCGCGCGCGGAGCCACGGCTTATGCAAGCAAGATCTCG 520
Qy 181 AGGAGCAGTTACGGTCTGTCGAGTGTGATGAACTCATGACTGTACTCTACCCAGAT 240
Db 521 AGGAGCAGTTACGGTCTGTCGAGTGTGATGAACTCATGACTGTACTCTACCCAGAT 580
Qy 241 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAACAG 300
Db 581 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAACAG 640
Qy 301 CCAACCTCAACTCAAGACAGAGAGACTATAAATTTGCTGACGACATATTATAACAG 360
Db 641 CCAACCTCAACTCAAGACAGAGAGACTATAAATTTGCTGACGACATATTATAACAG 700
Qy 361 AGATCTTTGAAAAATGATAAATGAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 420
Db 701 AGATCTTTGAAAAATGATAAATGAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 760
Qy 421 GTATAGATGTGGGAGAGAGTTGGAGTCCGCAAAACACCTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGTGGGAGAGAGTTGGAGTCCGCAAAACACCTCTTTAAACCTCCATGTG 820
Qy 481 TGTCCGCTACAGATGTGGGAGTTGCTCAATAGTGGGGGCTGCAGTGCATGAACACCA 540
Db 821 TGTCCGCTACAGATGTGGGAGTTGCTCAATAGTGGGGGCTGCAGTGCATGAACACCA 880
Qy 541 GCACGAGTACCTCAGCAAGACGTTATTGAAATTAAGTGCCTCTCTCTCAAGGCCCA 600
Db 881 GCACGAGTACCTCAGCAAGACGTTATTGAAATTAAGTGCCTCTCTCTCAAGGCCCA 940
Qy 601 AACAGTAAACATCAGTTTGGCAATCACACTTCTCCGATGCATGCTCTAACTCGAG 660
Db 941 AACAGTAAACATCAGTTTGGCAATCACACTTCTCCGATGCATGCTCTAACTCGAG 1000
Qy 661 TTTACAGCAAGTTTCAATTCATTATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720
Db 1001 TTTACAGCAAGTTTCAATTCATTATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 1060
Qy 721 AGGAGCGCAACAGACTGCCCCCAATTTACATGTGGAAATATCATCTGCAGATGCC 780
Db 1061 AGGAGCGCAACAGACTGCCCCCAATTTACATGTGGAAATATCATCTGCAGATGCC 1120
Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC 840
Db 1121 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC 1180
Qy 841 ATGACATCTGTGGACCAAAACAGAGCTGGAGTGAAGAGACCTGTCACTGTGTCTGCAGAG 900
Db 1181 ATGACATCTGTGGACCAAAACAGAGCTGGAGTGAAGAGACCTGTCACTGTGTCTGCAGAG 1240
Qy 901 CGGGCTTCGGCTCGCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1241 CGGGCTTCGGCTCGCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 1300
Qy 961 GTGTCTGTAAAAAACAACCTCTTCCCGAGCAATGTGGGCAACCCGAGAAATTTGATGAAA 1020
Db 1301 GTGTCTGTAAAAAACAACCTCTTCCCGAGCAATGTGGGCAACCCGAGAAATTTGATGAAA 1360
Qy 1021 ACACATGCGAGTGTATGTATTAAGAACTGCTCCCGAGAACTCAACCCCTAAATCTCGAA 1080
Db 1361 ACACATGCGAGTGTATGTATTAAGAACTGCTCCCGAGAACTCAACCCCTAAATCTCGAA 1420
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAAGAGTTC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAAGAGTTC 1480
Qy 1141 ACCACCAACATGCGAGCTTTACAGACGGCCATGTACAGACCGCCAGAGGCTTGTGAGC 1200

Db 1481 ACCACCAACATGCGAGCTTTACAGACGGCCATGTACAGACCGCCAGAGGCTTGTGAGC 1540
Qy 1201 CAGATTTTCATATAGTAGAAGAGTGTGCTGTTGTGCTCCCTTCATATTGGCAAGACAC 1260
Db 1541 CAGATTTTCATATAGTAGAAGAGTGTGCTGTTGTGCTCCCTTCATATTGGCAAGACAC 1600
Qy 1261 AAATGAGCTAAGATTTACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAAACTGTGT 1320
Db 1601 AAATGAGCTAAGATTTACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAAACTGTGT 1660
Qy 1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGACCCCTTGTGGGTCCATGCTAACAAAGACA 1380
Db 1661 TGCCACAGTAGAAGTGTCTGTGAACAGAGACCCCTTGTGGGTCCATGCTAACAAAGACA 1720
Qy 1381 AAAGTCTCTCTTCTGAAACCATGTGGATTAACATTTTACAGAAATGAGCTGGAGCTCATCTG 1440
Db 1721 AAAGTCTCTCTTCTGAAACCATGTGGATTAACATTTTACAGAAATGAGCTGGAGCTCATCTG 1780
Qy 1441 CAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCAAGATTTCTCTC 1500
Db 1781 CAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCAAGATTTCTCTC 1840
Qy 1501 TTGTGATTTCTTTAAAGAACTGATATATTTTATTTCCACTAAATAATTTGTTCTGC 1560
Db 1841 TTGTGATTTCTTTAAAGAACTGATATATTTTATTTCCACTAAATAATTTGTTCTGC 1900
Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACAAATTTGGTAAACACTCAGTGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATAAATGAATGAAATTTGTATT 1655
Db 1961 GCAAAATATGTTTAAATAAATGAATGAAATTTGTATT 1995

RESULT 11

US-09-355-700-7

Sequence 7, Application US/09355700

Patent No. 6361946

GENERAL INFORMATION:

APPLICANT: Ludwig Institute for Cancer Research

Heisinki University Licensing

Alitalo, Kari (U.S. only)

Joukov, Vladimir (U.S. only)

TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)

Protein and Gene, Mutants Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/355,700

FILING DATE: 05-NOV-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/795,430

FILING DATE: 05-FEB-1997

APPLICATION NUMBER: PCT/FI96/00427

FILING DATE: 01-AUG-1996

APPLICATION NUMBER: 08/671,573

FILING DATE: 28-JUN-1996

APPLICATION NUMBER: 08/601,132

FILING DATE: 14-FEB-1996

APPLICATION NUMBER: 08/585,895

```

; FILING DATE: 12-JAN-1996
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/34140
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1997 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1608
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-355-700-7

Query Match      98.7%; Score 1651.8; DB 4; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GTCCCTCCACATGCACTCCCTGGCTTCTCTCTGCTGGGCTGTTCTCTGCTCGCGCTG 60
DB      341  GTCCCTCCACATGCACTCCCTGGCTTCTCTCTGCTGGGCTGTTCTCTGCTCGCGCTG 400

QY      61  CGCTGCTCCGCGTCCCTCGGAGCGCCCGCGCGCGCGCGCTTCGAGTCCGGAATCG 120
DB      401  CGCTGCTCCGCGTCCCTCGGAGCGCCCGCGCGCGCGCTTCGAGTCCGGAATCG 460

QY      121  ACCTCTCGAGCGCGAGCCCGAGCGCGCGAGCGCGCTTATGACGAAAGATCTGG 180
DB      461  ACCTCTCGAGCGCGAGCCCGAGCGCGCGAGCGCGCTTATGACGAAAGATCTGG 520

QY      181  AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTCTACTTACCCAGAAT 240
DB      521  AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTCTACTTACCCAGAAT 580

QY      241  ATTGGAATAATGTAACAAGTGTGAGTGTAGTAAAGAAAGAGCGCTGGCAACATPACAGAGACAG 300
DB      581  ATTGGAATAATGTAACAAGTGTGAGTGTAGTAAAGAAAGAGCGCTGGCAACATPACAGAGACAG 640

QY      301  CCAACCTCACTCAGGACGAGAGACTATAAATTTGCTGAGACACATTATAATACAG 360
DB      641  CCAACCTCACTCAGGACGAGAGACTATAAATTTGCTGAGACACATTATAATACAG 700

QY      361  AGATCTTGAAGAAGTATTGATTAATCAGTGGAGAAAGACTCAATGATGCCCGGAGGTGT 420
DB      701  AGATCTTGAAGAAGTATTGATTAATCAGTGGAGAAAGACTCAATGATGCCCGGAGGTGT 760

QY      421  GTATAGATGCGGGAAGAGTTTGGAGTCCGCAACAAACCTTCTTTAAACCTCCATGTG 480
DB      761  GTATAGATGCGGGAAGAGTTTGGAGTCCGCAACAAACCTTCTTTAAACCTCCATGTG 820

QY      481  TGTCCGCTTACAGATGTGGGGGTTGCTGCAATAGTGGAGGGGCTGCAGTGCATGAACACCA 540
DB      821  TGTCCGCTTACAGATGTGGGGGTTGCTGCAATAGTGGAGGGGCTGCAGTGCATGAACACCA 880

QY      541  GCACGAGTACTCAGCAAGAGCTTATTGAAATACAGTGCCTCTCTCAGGCGCCCA 600
DB      881  GCACGAGTACTCAGCAAGAGCTTATTGAAATACAGTGCCTCTCTCAGGCGCCCA 940

QY      601  AACAGTAAACATCAGTTTTCCTCAATCAGCTTCTGCGGATGCATGCTCAAACTGGATG 660
DB      941  AACAGTAAACATCAGTTTTCCTCAATCAGCTTCTGCGGATGCATGCTCAAACTGGATG 1000

```

```

QY      661  TTTACAGACAGAGTTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACACAGTGTG 720
DB      1001  TTTACAGACAGAGTTTCATTCCATTATTAGACGTTCCCTGCCAGCAACTACACAGTGTG 1060

QY      721  AGGAGCGAACAAGACCTGCCACCAATACATGTGGATAATACATCTGCGAGATGCC 780
DB      1061  AGGAGCGAACAAGACCTGCCACCAATACATGTGGATAATACATCTGCGAGATGCC 1120

QY      781  TGGCTCAGGAAGATTTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTC 840
DB      1121  TGGCTCAGGAAGATTTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTC 1180

QY      841  ATGACATCTGTGGACCAAAAACAGAGCTGGATGAAGAGACCTGTGCTGCTGCGAGAG 900
DB      1181  ATGACATCTGTGGACCAAAAACAGAGCTGGATGAAGAGACCTGTGCTGCTGCGAGAG 1240

QY      901  CGGGCTTTCGGCTGCGAGCTGTGGACCCCAAAAGAACTAGACAGAGAACTCATGCCAGT 960
DB      1241  CGGGCTTTCGGCTGCGAGCTGTGGACCCCAAAAGAACTAGACAGAGAACTCATGCCAGT 1300

QY      961  GTGCTGTAAAAACAAACTCTTCCCAGCCCAATGTGGGGCCCAACCGAGAAATTTGATGAA 1020
DB      1301  GTGCTGTAAAAACAAACTCTTCCCAGCCCAATGTGGGGCCCAACCGAGAAATTTGATGAA 1360

QY      1021  ACACATGCCAGTGTGTATGTAAGAAAGAACTGCCCCAGAGAAATCAACCCCTAAATCCTGAA 1080
DB      1361  ACACATGCCAGTGTGTATGTAAGAAAGAACTGCCCCAGAGAAATCAACCCCTAAATCCTGAA 1420

QY      1081  AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTCC 1140
DB      1421  AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTCC 1480

QY      1141  ACCACAAACATGACGCTGTGTACAGCGCCATGTAGAACCGCCAGAGCGCTGTGAGC 1200
DB      1481  ACCACAAACATGACGCTGTGTACAGCGCCATGTAGAACCGCCAGAGCGCTGTGAGC 1540

QY      1201  CAGGATTTTCATATAGTGAAGAGTGTGCTGCTGTGCTCCCTTCATATTTGGCAAGACAC 1260
DB      1541  CAGGATTTTCATATAGTGAAGAGTGTGCTGCTGTGCTCCCTTCATATTTGGCAAGACAC 1600

QY      1261  AATGAGCTAGATTTGTAAGTCTGTTTCCAGTTCATCGATTTTCTATTATGGAAGAACTGT 1320
DB      1601  AATGAGCTAGATTTGTAAGTCTGTTTCCAGTTCATCGATTTTCTATTATGGAAGAACTGT 1660

QY      1321  TGCCACAGTGTAGAACTGTGTGAAACAGAGAGACCTGTGGGTCCATGCTTAACAAAGACA 1380
DB      1661  TGCCACAGTGTAGAACTGTGTGAAACAGAGAGACCTGTGGGTCCATGCTTAACAAAGACA 1720

QY      1381  AAAAGTCTGCTTTTCTGAACCATGTGGATACTTTTACAGAAATGGAGCTGGAGCTCATCTG 1440
DB      1721  AAAAGTCTGCTTTTCTGAACCATGTGGATACTTTTACAGAAATGGAGCTGGAGCTCATCTG 1780

QY      1441  CAAAAGGCTCTGTGTAAAGACTGGTTTTCGCAATGACCAACAGCAAGATTTTCTCTC 1500
DB      1781  CAAAAGGCTCTGTGTAAAGACTGGTTTTCGCAATGACCAACAGCAAGATTTTCTCTC 1840

QY      1501  TTGTGATTTCTTTAAAGAAAGTATATAATTTATTTCCACTAAAAATATTTGTTTCTGC 1560
DB      1841  TTGTGATTTCTTTAAAGAAAGTATATAATTTATTTCCACTAAAAATATTTGTTTCTGC 1900

QY      1561  ATTCAATTTTATAGCAACAACTTTGTAAGAAAGTGTGTAAGTGTGATCAATTTTATATCAT 1620
DB      1901  ATTCAATTTTATAGCAACAACTTTGTAAGAAAGTGTGTAAGTGTGATCAATTTTATATCAT 1960

QY      1621  GCAAAATATGTTTAAAAATAAAATGAAATTTGATT 1655
DB      1961  GCAAAATATGTTTAAAAATAAAATGAAATTTGATT 1995

```

RESULT 12
 US-08-601-132-32
 ; Sequence 32, Application US/08601132

Db 1781 CAAAAGGCTCTGTAAAGATGGTTTCTGCCAATGACCAACAGCAAGATTTCTCTC 1840
 Qy 1501 TTGTGATTTCTTTAAAGAAAGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
 Db 1841 TTGTGATTTCTTTAAAGAAAGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1900
 Qy 1561 ATTCAATTTTATAGCAACAAATGTTGTAAGAACTCAGTGATCAATATTTTATATCAT 1620
 Db 1901 ATTCAATTTTATAGCAACAAATGTTGTAAGAACTCAGTGATCAATATTTTATATCAT 1960
 Qy 1621 GCAAAATATGTTTAAATAAATAAATGAATTTGTTAT 1655
 Db 1961 GCAAAATATGTTTAAATAAATAAATGAATTTGTTAT 1995

RESULT 13
 US-08-671-573B-32
 ; Sequence 32, Application US/08671573B
 ; Patent No. 6645933
 ; GENERAL INFORMATION:
 ; APPLICANT: Alitalo, Kari
 ; APPLICANT: Joukov, Vladimir
 ; TITLE OF INVENTION: Receptor Ligand
 ; NUMBER OF SEQUENCES: 58
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/671,573B
 ; FILING DATE: 28-JUN-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/601,132
 ; FILING DATE: 14-FEB-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/585,895
 ; FILING DATE: 12-JAN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/510,133
 ; FILING DATE: 01-AUG-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gass, David A.
 ; REGISTRATION NUMBER: 38,153
 ; REFERENCE/DOCKET NUMBER: 28967/33348
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1997 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 352..1608
 ; US-08-671-573B-32

Query Match 98.7%; Score 1651.8; DB 4; Length 1997;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCACTCGCTGGGTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 60
 Db 341 GTCTTCCACCATGCACTTGTGGGTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 400
 Qy 61 CGCTGCTCCGGGTCTCGGAGGGCCCGCGCGCGCGCGCTTCGAGTTCGGACTCG 120
 Db 401 CGCTGCTCCGGGTCTCGGAGGGCCCGCGCGCGCGCGCTTCGAGTTCGGACTCG 460
 Qy 121 ACCTCTCGGACGCGGAGCCCGGCGGCGGCGGCGGCGGCTTATGCAAGCAAGATCTGG 180
 Db 461 ACCTCTCGGACGCGGAGCCCGGCGGCGGCGGCGGCTTATGCAAGCAAGATCTGG 520
 Qy 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAAT 240
 Db 521 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAAT 580
 Qy 241 ATTGGAAAATGTACAGTGTCAAGTGAAGGAAAGGAGGTGGCAACATTAACAGAACAGG 300
 Db 581 ATTGGAAAATGTACAGTGTCAAGTGAAGGAAAGGAGGTGGCAACATTAACAGAACAGG 640
 Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATATAATTTGCTCGACACATTAATAACAG 360
 Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATATAATTTGCTCGACACATTAATAACAG 700
 Qy 361 AGATCTTGAAGATATTGATATGATGTGAGAGAGACTCAATGCATGCCACGGAGGTCT 420
 Db 701 AGATCTTGAAGATATTGATATGATGTGAGAGAGACTCAATGCATGCCACGGAGGTCT 760
 Qy 421 GTATAGATGTGGGAGGAGTGTGGAGTGCACACAAACACCTTTCTTAAACCTCCATGTG 480
 Db 761 GTATAGATGTGGGAGGAGTGTGGAGTGCACACAAACACCTTTCTTAAACCTCCATGTG 820
 Qy 481 TGTCCGTCTACAGATGTGGGGTGTGTCGAATAGTAGAGGGGCTGCAGTGCATGAACACCA 540
 Db 821 TGTCCGTCTACAGATGTGGGGTGTGTCGAATAGTAGAGGGGCTGCAGTGCATGAACACCA 880
 Qy 541 GCACGAGCTACCTCAGCAAGAGCTTATTTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
 Db 881 GCACGAGCTACCTCAGCAAGAGCTTATTTGAAATTTACAGTGCCTCTCTCAAGGCCCA 940
 Qy 601 AACAGTAAATCAGTTTGTGCAATCAACCTTCTGCGGATGCATGTCTAACTGGATG 660
 Db 941 AACAGTAAATCAGTTTGTGCAATCAACCTTCTGCGGATGCATGTCTAACTGGATG 1000
 Qy 661 TTTACAGACAAAGTTCATTTCCATTTAGACGTTTCCCTGCCAGCAACACTACCAAGTGC 720
 Db 1001 TTTACAGACAAAGTTCATTTCCATTTAGACGTTTCCCTGCCAGCAACACTACCAAGTGC 1060
 Qy 721 AGGAGCAACAGACCTGCCCCCAGCAATTTACATGTGGAATTAATCACATCTCCAGATGCC 780
 Db 1061 AGGAGCAACAGACCTGCCCCCAGCAATTTACATGTGGAATTAATCACATCTCCAGATGCC 1120
 Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGTGCATCAACAGATGATTC 840
 Db 1121 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGTGCATCAACAGATGATTC 1180
 Qy 841 ATGACATCTGTGGACCAAAACAGGAGCTGGATGAAGAGACCTGTGTGAGTGTCTGCAGAG 900
 Db 1181 ATGACATCTGTGGACCAAAACAGGAGCTGGATGAAGAGACCTGTGTGAGTGTCTGCAGAG 1240
 Qy 901 CGGGGCTTCGGCTTCAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
 Db 1241 CGGGGCTTCGGCTTCAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 1300
 Qy 961 GTGCTGTGTAATAAACAACCTTTTCCCGAGCAATGTGGGGCCCAACGAGAAATTTGATGAA 1020
 Db 1301 GTGCTGTGTAATAAACAACCTTTTCCCGAGCAATGTGGGGCCCAACGAGAAATTTGATGAA 1360
 Qy 1021 ACATATGCGAGTGTGTATATAAAGAACTGCGCCAGAAATCAACCCCTAAATTCCTGGAA 1080
 Db 1361 ACATATGCGAGTGTGTATATAAAGAACTGCGCCAGAAATCAACCCCTAAATTCCTGGAA 1420

QY 1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
Db 1421 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1480
QY 1141 ACCACCAAAATCAGCTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1200
Db 1481 ACCACCAAAATCAGCTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1540
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTCCCTTCATATGGCCAAAGACAC 1260
Db 1541 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTCCCTTCATATGGCCAAAGACAC 1600
QY 1261 AATGAGCTAAGATTTACTGTTTCCAGTTCATGATTTCTATATGGAAGAGTGT 1320
Db 1601 AATGAGCTAAGATTTACTGTTTCCAGTTCATGATTTCTATATGGAAGAGTGT 1660
QY 1321 TGCCACAGTAGAACTGTGTGAACAGAGAGACCCCTTGTGGTCCATGTCTAAACAAAGACA 1380
Db 1661 TGCCACAGTAGAACTGTGTGAACAGAGAGACCCCTTGTGGTCCATGTCTAAACAAAGACA 1720
QY 1381 AAGTCTGCTTTCCTGAACCATGTGATAACTTTACAGAAATGACTGGAGCTCATCTG 1440
Db 1721 AAGTCTGCTTTCCTGAACCATGTGATAACTTTACAGAAATGACTGGAGCTCATCTG 1780
QY 1441 CAAAGGCTCTTGTAAAGAGTGTGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1500
Db 1781 CAAAGGCTCTTGTAAAGAGTGTGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1840
QY 1501 TTGTGATTTCTTAAAGAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Db 1841 TTGTGATTTCTTAAAGAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1900
QY 1561 ATTCAATTTTATAGCAACCAATTTGTTAACTCACTGTGATCAATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACCAATTTGTTAACTCACTGTGATCAATTTTATATCAT 1960
QY 1621 GCAAAATATGTTTAAATAAATGAAATTTGTAAT 1655
Db 1961 GCAAAATATGTTTAAATAAATGAAATTTGTAAT 1995

RESULT 14
US-08-999-811-3
; Sequence 3, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:
; APPLICANT: HU JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMEUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,811
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1526 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 71..142
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 143..1120
FEATURE:
NAME/KEY: CDS
LOCATION: 71..1120
US-08-999-811-3
Query Match 91.2%; Score 1526; DB 2; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 CGAGGCCACGGCTTATGCAAGCAAGATCTGGAGGAGCAGTTACGGTCTGTGTCAGTGT 208
Db 1 CGAGGCCACGGCTTATGCAAGCAAGATCTGGAGGAGCAGTTACGGTCTGTGTCAGTGT 60
QY 209 AGATGAATCATGACTCTACTCTACCCAGATAATTTGAAATATGTAAGTGTGAGTAAAG 268
Db 61 AGATGAATCATGACTCTACTCTACCCAGATAATTTGAAATATGTAAGTGTGAGTAAAG 120
QY 269 GAAAGGAGGCTGGCAACATTAACAGAGAACAGGCCAACCTCAACTCAAGGAGCAGAGAGAC 328
Db 121 GAAAGGAGGCTGGCAACATTAACAGAGAACAGGCCAACCTCAACTCAAGGAGCAGAGAGAC 180
QY 329 TATAAATTTTCTGCAGCAGCATTTATATACAGAGATCTTGAAAGATTTGATATGAGTG 388
Db 181 TATAAATTTTCTGCAGCAGCATTTATATACAGAGATCTTGAAAGATTTGATATGAGTG 240
QY 389 GAGAAAGACTCAATGCAATGCCACGGAGGCTGTGTATAGATGTGGGAGAGAGTTTGGAGT 448
Db 241 GAGAAAGACTCAATGCAATGCCACGGAGGCTGTGTATAGATGTGGGAGAGAGTTTGGAGT 300
QY 449 CGCAGCAAAACACCTTCTTTAAACCTCCATGTGTGTCGCTACAGATGTGGGGTGTGCTG 508
Db 301 CGCAGCAAAACACCTTCTTTAAACCTCCATGTGTGTCGCTACAGATGTGGGGTGTGCTG 360
QY 509 CAATAGTAGGGGCTGCAGTGTGATGAACACAGCAGCAGCTACTCTCAGCAAGAGCTTATT 568
Db 361 CAATAGTAGGGGCTGCAGTGTGATGAACACAGCAGCAGCTACTCTCAGCAAGAGCTTATT 420
QY 569 TGAATTTACAGTGCCTCTCTCTCAAGGCCCAACCAAGTAAACATCAGTTTGGCCAAATCA 628
Db 421 TGAATTTACAGTGCCTCTCTCTCAAGGCCCAACCAAGTAAACATCAGTTTGGCCAAATCA 480
QY 629 CACTTCTCTCCGATGCAATGCTTAACTGATGTTTACAGACAGTTCATTCCATTATTAG 688
Db 481 CACTTCTCTCCGATGCAATGCTTAACTGATGTTTACAGACAGTTCATTCCATTATTAG 540
QY 689 ACGTTCCCTGCCAGCAACACTTACCACAGTGTGAGGAGCGAAACAAAGACCTGCCCCACAA 748
Db 541 ACGTTCCCTGCCAGCAACACTTACCACAGTGTGAGGAGCGAAACAAAGACCTGCCCCACAA 600
QY 749 TTACATGTGGAATTAATCAGATCTGCAGATGCTGGCTCAGAGAGATTTATGTTTCTC 808
Db 601 TTACATGTGGAATTAATCAGATCTGCAGATGCTGGCTCAGAGAGATTTATGTTTCTC 660

```

QY 809 GGATGCTGGAGATGATCTCAACAGATGGATTCCATGACATCTGTGGACCAAAACAGAGAGCT 868
Db 661 GGATGCTGGAGATGATCTCAACAGATGGATTCCATGACATCTGTGGACCAAAACAGAGAGCT 720
QY 869 GGATGAGAGAGACCTGTAGTGTGTCTGAGAGCGGGGCTTCGGCTCGCAGCTGTGGACC 928
Db 721 GGATGAGAGAGACCTGTAGTGTGTCTGAGAGCGGGGCTTCGGCTCGCAGCTGTGGACC 780
QY 929 CCACAAAGAACTAGACAGAAACTCATGCGCATGTGTCTGTAAACAAACAACTCTTCCCCAG 988
Db 781 CCACAAAGAACTAGACAGAAACTCATGCGCATGTGTCTGTAAACAAACAACTCTTCCCCAG 840
QY 989 CCAATGTGGGCCCAACCGAGAAATTTGATGAACACACATGCCAGTGTGTATGAAGAAAC 1048
Db 841 CCAATGTGGGCCCAACCGAGAAATTTGATGAACACACATGCCAGTGTGTATGAAGAAAC 900
QY 1049 CTGCCCCAGAAATCAACCCCTAAATCTCTGGAAATGTCCCTGTGAATGTACAGAAAGTCC 1108
Db 901 CTGCCCCAGAAATCAACCCCTAAATCTCTGGAAATGTCCCTGTGAATGTACAGAAAGTCC 960
QY 1109 ACAGAAATGCTGTGTAAGGAAGAGAGTTCACACCAACACATGCAGCTGTACAGACG 1168
Db 961 ACAGAAATGCTGTGTAAGGAAGAGAGTTCACACCAACACATGCAGCTGTACAGACG 1020
QY 1169 GCCATGTACGAACCGCCAGAGAGCTTGTGAGCCAGGATTTTCATATAGTGAAGAGTGTG 1228
Db 1021 GCCATGTACGAACCGCCAGAGAGCTTGTGAGCCAGGATTTTCATATAGTGAAGAGTGTG 1080
QY 1229 TCGTGTGTCCCTTCATATGCGAAAGACACAAATAGAGTGTACTGTGTTCCTCA 1288
Db 1081 TCGTGTGTCCCTTCATATGCGAAAGACACAAATAGAGTGTACTGTGTTCCTCA 1140
QY 1289 GTTCATCGAATTTCTATTATGAAACCTGTGTGCCAGTAGAACTGTCTGTGAACAGA 1348
Db 1141 GTTCATCGAATTTCTATTATGAAACCTGTGTGCCAGTAGAACTGTCTGTGAACAGA 1200
QY 1349 GAGACCTTGTGGTCCATGTCTACAAAGACAAAGAGTGTCTTCTCTGAACCATGTGGA 1408
Db 1201 GAGACCTTGTGGTCCATGTCTACAAAGACAAAGAGTGTCTTCTCTGAACCATGTGGA 1260
QY 1409 TAATTTACAGAAATGAGTGGAGTCTCATCTGAAAGGCTCTGTAAAGAGCTGGTTTT 1468
Db 1261 TAATTTACAGAAATGAGTGGAGTCTCATCTGAAAGGCTCTGTAAAGAGCTGGTTTT 1320
QY 1469 CTGCAATGACCAACAGCCAGAGATTTTCCCTCTGTGATTTCTTAAAGAACTACTATA 1528
Db 1321 CTGCAATGACCAACAGCCAGAGATTTTCCCTCTGTGATTTCTTAAAGAACTACTATA 1380
QY 1529 TAATTTATTTCCACTAAATAATTTGTTTCTGCATTCATTTTATAGCAACAACTATGGT 1588
Db 1381 TAATTTATTTCCACTAAATAATTTGTTTCTGCATTCATTTTATAGCAACAACTATGGT 1440
QY 1589 AAACTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATAAATGAAAA 1648
Db 1441 AAACTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATAAATGAAAA 1500
QY 1649 TTGATTTTAAAAAATAAATAAATAA 1674
Db 1501 TTGATTTTAAAAAATAAATAAATAA 1526

```

```

RESULT 15
US-09-042-105-3
; Sequence 3, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX

```

```

; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: HERETHWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TO BE ASSIGNED
; FILING DATE: 24-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ERIC K. STEFFE
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 71..142
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 143..1120
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 71..1120
; US-09-042-105-3

```

Query Match 91.2%; Score 1526; DB 3; Length 1526;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 149 CGAGCCACCGCTTATGCAAGCAAGATCTGGAGGAGCAGTTACGCTGTGTCAGGTGT 208
Db 1 CGAGCCACCGCTTATGCAAGCAAGATCTGGAGGAGCAGTTACGCTGTGTCAGGTGT 60
QY 209 AGATGAACCTCATGACTGTACTCTACCCAGAAATATTGAAAAATGTACAAAGTGTCAAG 268
Db 61 AGATGAACCTCATGACTGTACTCTACCCAGAAATATTGAAAAATGTACAAAGTGTCAAG 120
QY 269 GAAAGAGGCTGGCAACATACAGAGAACAGCCCACTCACTCAAGGACAGAGAGAC 328
Db 121 GAAAGAGGCTGGCAACATACAGAGAACAGCCCACTCACTCAAGGACAGAGAGAC 180
QY 329 TATAAAATTTGCTGAGCAGACATTATAATACAGAGATCTTGAAGATTTGATAATGAGTG 388
Db 181 TATAAAATTTGCTGAGCAGACATTATAATACAGAGATCTTGAAGATTTGATAATGAGTG 240
QY 389 GAGAAAGACTCAATGCATGCCACCGGAGTGTGTATAGATGTGGGAGGAGTTGGAGT 448

```

Db 241 GAGAAAGACTCAATGCATGCCACGGGAGGTGTGTATAGATGTGGGAGGAGTTGGAGT 300
 Qy 449 CGGACAAACACCTCTTTAAACCTCCATGTGTGTCCGTCTACAGATGTGGGGTGTGCTG 508
 Db 301 CGGACAAACACCTCTTTAAACCTCCATGTGTGTCCGTCTACAGATGTGGGGTGTGCTG 360
 Qy 509 CAATAGTGAAGGGGTGAGTGCATGAACACACGACGAGCTACCTCAGCAAGAGCTTTATT 568
 Db 361 CAATAGTGAAGGGGTGAGTGCATGAACACACGACGAGCTACCTCAGCAAGAGCTTTATT 420
 Qy 569 TGAATTTACGTGCTCTCTCTAAGGCCCCAAACCAAGTAAACATCAATCTTTGCCAATCA 628
 Db 421 TGAATTTACGTGCTCTCTCTAAGGCCCCAAACCAAGTAAACATCAATCTTTGCCAATCA 480
 Qy 629 CACTTCCTGCGATGCTCTTAACCTGAGTGTATACAGACAGTTCATTTCCATTTATAG 688
 Db 481 CACTTCCTGCGATGCTCTTAACCTGAGTGTATACAGACAGTTCATTTCCATTTATAG 540
 Qy 689 ACGTTCCTGCCAGCAACACTACACAGTGTACGACGAGCAAGACCTGCCCCACCAA 748
 Db 541 ACGTTCCTGCCAGCAACACTACACAGTGTACGACGAGCAAGACCTGCCCCACCAA 600
 Qy 749 TTACATGTGGATTAATCACATCTCAGATGCTGGCTCAGGAAGATTTATGTTTCCTC 808
 Db 601 TTACATGTGGATTAATCACATCTCAGATGCTGGCTCAGGAAGATTTATGTTTCCTC 660
 Qy 809 GGATGCTGGAGATCACTCAACAGATGGATTCATGACATCTGTGGACCAACCAAGGAGCT 868
 Db 661 GGATGCTGGAGATCACTCAACAGATGGATTCATGACATCTGTGGACCAACCAAGGAGCT 720
 Qy 869 GGATGAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTGCCAGCTGTGGACC 928
 Db 721 GGATGAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTGCCAGCTGTGGACC 780
 Qy 929 CCACAAAGAACTAGACAGAACTCATCCAGTGTCTGTAAACCAACAACTCTTCCCCAG 988
 Db 781 CCACAAAGAACTAGACAGAACTCATCCAGTGTCTGTAAACCAACAACTCTTCCCCAG 840
 Qy 989 CCAATGTGGGCCCAACCGAGAAATTTGATGAACACATGCCAGTGTGTATGTAAAGAAC 1048
 Db 841 CCAATGTGGGCCCAACCGAGAAATTTGATGAACACATGCCAGTGTGTATGTAAAGAAC 900
 Qy 1049 CTGCCCCAGAAATCAACCCCTAAATCTGGAAATGTCCCTGTAATGTACAGAAAGTCC 1108
 Db 901 CTGCCCCAGAAATCAACCCCTAAATCTGGAAATGTCCCTGTAATGTACAGAAAGTCC 960
 Qy 1109 ACAGAAATGCTGTGTAAAGAAAGAAAGTTCCACCAACCAACATGCAGCTGTGTACAGACG 1168
 Db 961 ACAGAAATGCTGTGTAAAGAAAGAAAGTTCCACCAACCAACATGCAGCTGTGTACAGACG 1020
 Qy 1169 GCCATGTACGAAACCGCCAGAGGCTTGTGAGCCAGGATTTTCATATAGTGAAGAGTGTG 1228
 Db 1021 GCCATGTACGAAACCGCCAGAGGCTTGTGAGCCAGGATTTTCATATAGTGAAGAGTGTG 1080
 Qy 1229 TCGTGTGTCCCTTCATATGGCAAGACCAACATGAGCTAAGATGTGTACTGTGTTCCTCA 1288
 Db 1081 TCGTGTGTCCCTTCATATGGCAAGACCAACATGAGCTAAGATGTGTACTGTGTTCCTCA 1140
 Qy 1289 GTTCATGCAATTTCTATATGAAACCTGTGTGCCACAGTAGAACTGTCTGTGAACAGA 1348
 Db 1141 GTTCATGCAATTTCTATATGAAACCTGTGTGCCACAGTAGAACTGTCTGTGAACAGA 1200
 Qy 1349 GAGACCTGTGGGTCCATGTAACAAAGACAAAGTCTGCTTTCTGGAACCATGTGGA 1408
 Db 1201 GAGACCTGTGGGTCCATGTAACAAAGACAAAGTCTGCTTTCTGGAACCATGTGGA 1260
 Qy 1409 TAACTTTACAGAAATGAGCTGAGCTCATCTGCAAAAGGCTCTTGTAAAGACTGGTTTT 1468
 Db 1261 TAACTTTACAGAAATGAGCTGAGCTCATCTGCAAAAGGCTCTTGTAAAGACTGGTTTT 1320
 Qy 1469 CTGCCAATGACCAACGACCAAGATTTTCTCTGTGATTTCTTTAAAGAAATGACTATA 1528
 Db 1321 CTGCCAATGACCAACGACCAAGATTTTCTCTGTGATTTCTTTAAAGAAATGACTATA 1380

Qy 1529 TAAATTTATTCACCTAAATAATTTCTTCTGCAATTCATTTTATAGCAACCAACATTCGT 1588
 Db 1381 TAAATTTATTCACCTAAATAATTTCTTCTGCAATTCATTTTATAGCAACCAACATTCGT 1440
 Qy 1589 AAAAATCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATATAAATGAAAA 1648
 Db 1441 AAAAATCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATATAAATGAAAA 1500
 Qy 1649 TTGTATTTATAAAAAAATAAAAAA 1674
 Db 1501 TTGTATTTATAAAAAAATAAAAAA 1526

Search completed: February 25, 2004, 22:07:19
 Job time : 152 secs

[illegible]

Db 61 CGCTGCTCCGGGTCCTCGAGCGCCCGCCGCGCGCTTCGAGTCGGACTCG 120
Qy 121 ACCTCTCGAGCGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGAATCTGG 180
Db 121 ACCTCTCGAGCGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGAATCTGG 180
Qy 181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 240
Db 181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 240
Qy 241 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAACAG 300
Db 241 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAACAG 300
Qy 301 CCAACCTCACTCAAGGACAGAGAGACTATAAAATTTGCTGAGACACATTAATATACAG 360
Db 301 CCAACCTCACTCAAGGACAGAGAGACTATAAAATTTGCTGAGACACATTAATATACAG 360
Qy 361 AGATCTTTGAAAAATGATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
Db 361 AGATCTTTGAAAAATGATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
Qy 421 GTATAGATGTGGGAAGAGATTGGAGTCCGCAACAAACACCTTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGAAGAGATTGGAGTCCGCAACAAACACCTTTTAAACCTCCATGTG 480
Qy 481 TGTCCGCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGCTGCAATGCAACCA 540
Db 481 TGTCCGCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGCTGCAATGCAACCA 540
Qy 541 GCACGAGTACCTCAGCAAGACGTTATTTGAAATTAACAGTGCCCTCTCTCAAGGCCCA 600
Db 541 GCACGAGTACCTCAGCAAGACGTTATTTGAAATTAACAGTGCCCTCTCTCAAGGCCCA 600
Qy 601 ARCCAGTAAACATCAGTTTCCCAATCACACTTCTCCGCGATCATGTCTAACTGGATG 660
Db 601 ARCCAGTAAACATCAGTTTCCCAATCACACTTCTCCGCGATCATGTCTAACTGGATG 660
Qy 661 TTTACAGCAAGTTCATTCCATTTATAGAGTTCCTCCGAGCAACACTACCAAGTGTG 720
Db 661 TTTACAGCAAGTTCATTCCATTTATAGAGTTCCTCCGAGCAACACTACCAAGTGTG 720
Qy 721 AGGAGCGAACAAGCTGCGCCACCAATACATGTGGAAATATCATCTGCGAGATGCC 780
Db 721 AGGAGCGAACAAGCTGCGCCACCAATACATGTGGAAATATCATCTGCGAGATGCC 780
Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 840
Db 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 840
Qy 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGTCTGATGTGTGAGAG 900
Db 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGTCTGATGTGTGAGAG 900
Qy 901 CGGGGCTTCGGCTCGCAGCTGTGGACCCACAAAGAACTAGACAGAACTATGCCACT 960
Db 901 CGGGGCTTCGGCTCGCAGCTGTGGACCCACAAAGAACTAGACAGAACTATGCCACT 960
Qy 961 GTGTCTGTAAAAACAATACTCTTCCAGCCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
Db 961 GTGTCTGTAAAAACAATACTCTTCCAGCCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
Qy 1021 ACATATGCGAGTGTATGTATTAAGAACTGCGCCCGAGAAATCAACCTTAATCTCGGAA 1080
Db 1021 ACATATGCGAGTGTATGTATTAAGAACTGCGCCCGAGAAATCAACCTTAATCTCGGAA 1080
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTTAAAGGAAAGAACTTC 1140
Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTTAAAGGAAAGAACTTC 1140
Qy 1141 ACCACCAAAATGAGCTGTATTAAGAGCGGCAATGTACGAAACCGCCAGAGGCTTGTGAGC 1200

Db 1141 ACCACCAAAATGAGCTGTATTAAGAGCGGCAATGTACGAAACCGCAGAGCTTGTGAGC 1200
Qy 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATATGGCAAGACAC 1260
Db 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATATGGCAAGACAC 1260
Qy 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAAACTGTGT 1320
Db 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAAACTGTGT 1320
Qy 1321 TGCCACAGTAGAAGTGTGTGAAACAGAGAGACCTTTGTGGGTCCATGCTTAACAAAGACA 1380
Db 1321 TGCCACAGTAGAAGTGTGTGAAACAGAGAGACCTTTGTGGGTCCATGCTTAACAAAGACA 1380
Qy 1381 AAAGTCTGTCTTCTGNAACCATGGGATTAACATTTTACAGAAATGAGCTGGAGCTCATCTG 1440
Db 1381 AAAGTCTGTCTTCTGNAACCATGGGATTAACATTTTACAGAAATGAGCTGGAGCTCATCTG 1440
Qy 1441 CAAAGGCGCTCTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1500
Db 1441 CAAAGGCGCTCTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1500
Qy 1501 TTGTGATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1560
Db 1501 TTGTGATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1560
Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Qy 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTATAAAAAA 1674

RESULT 2

US-09-921-143-1
; Sequence 1, Application US/09921143
; Publication No. US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI112P6
; CURRENT APPLICATION NUMBER: US/09/921,143
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-1

Query Match 100.0%; Score 1674; DB 11; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCTTCACCATGCATCGCTGGGCTTCTCTGTGGCGTGTCTCTGTCTGCCCGCTG 60
Db 1 GTCTTCACCATGCATCGCTGGGCTTCTCTGTGGCGTGTCTCTGTCTGCCCGCTG 60
Qy 61 CGCTGCTCCGGGTCTTCGAGAGCGCGCGCGCGCGCGCTTCGAGTCCGGACTCG 120
Db 61 CGCTGCTCCGGGTCTTCGAGAGCGCGCGCGCGCGCGCTTCGAGTCCGGACTCG 120
Qy 121 ACCTCTCGAGCGCGAGCCCGACCGGGCGAGGCCAGGGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGAGCGCGAGCCCGACCGGGCGAGGCCAGGGCTTATGCAAGCAAGATCTGG 180
Qy 181 AGGAGCAGTTACGGTGTGTGTCTGATGATGATGATGATGATGATGATGATGATGAT 240

181	AGGAGCAGTTACGGTCTGTGTCACGTGTAGATGAACATCATGACTGTACTCTTACCCAGAA	240
241	ATTGGAATAATGACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAAACAGG	300
241	ATTGGAATAATGACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAAACAGG	300
301	CCACCTCAACTCAAGGACGAAGAGACTATAAAATTTGCTGCAGCACATTATAATACAG	360
301	CCACCTCAACTCAAGGACGAAGAGACTATAAAATTTGCTGCAGCACATTATAATACAG	360
361	AGATCTTTGAAAGTATTGATAATGAGTCGAGAAAGACTCAATGATGCGACGGAGGTGT	420
361	AGATCTTTGAAAGTATTGATAATGAGTCGAGAAAGACTCAATGATGCGACGGAGGTGT	420
421	GTAATAGATGTTGGGAGAGATTTGGAGTCGGCAAAACACACTTCTTTAAACCTCATGTG	480
421	GTAATAGATGTTGGGAGAGATTTGGAGTCGGCAAAACACACTTCTTTAAACCTCATGTG	480
481	TGTCGGTCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGCTGCAGTCATGAACACCA	540
481	TGTCGGTCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGCTGCAGTCATGAACACCA	540
541	GCACGAGTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
541	GCACGAGTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
601	AACCAAGTAACAATCAGTTTTCGCAATCACACTTCTGCGGATGCATGCTAAATCTGATG	660
601	AACCAAGTAACAATCAGTTTTCGCAATCACACTTCTGCGGATGCATGCTAAATCTGATG	660
661	TTTACAGACAAATTCATTTCCATTTAGACGTTCCCTGCCAGCAACACTACCAAGTGC	720
661	TTTACAGACAAATTCATTTCCATTTAGACGTTCCCTGCCAGCAACACTACCAAGTGC	720
721	AGGAGCGGACAGACCTGCCACCAATTTACATGTGGAATTAATCAATCTGCAGATGCC	780
721	AGGAGCGGACAGACCTGCCACCAATTTACATGTGGAATTAATCAATCTGCAGATGCC	780
781	TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGATCC	840
781	TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGATCC	840
841	ATGACATCTGTGGACCAAAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGCAGAG	900
841	ATGACATCTGTGGACCAAAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGCAGAG	900
901	CGGGCTTCGGCTGCCAGCTGTGACCCCAAGAACTAGACAGAAACTCATGCCAGT	960
901	CGGGCTTCGGCTGCCAGCTGTGACCCCAAGAACTAGACAGAAACTCATGCCAGT	960
961	GTGCTGTAAAAAATACTTTCTCCAGCCCAATGTGGGGCCAAACCGAGAAATTTGATGAA	1020
961	GTGCTGTAAAAAATACTTTCTCCAGCCCAATGTGGGGCCAAACCGAGAAATTTGATGAA	1020
1021	ACATCCGCTGTGATGTAAAGAACCTTCCCGAGAACTCAACCCCTTAATCTCGAA	1080
1021	ACATCCGCTGTGATGTAAAGAACCTTCCCGAGAACTCAACCCCTTAATCTCGAA	1080
1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGATTTCC	1140
1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGATTTCC	1140
1141	ACCACCAACATGACGCTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC	1200
1141	ACCACCAACATGACGCTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC	1200
1201	CAGGATTTTCATATAGTGAAGAGTGTCTGTGTGTCCTTTCATATTTGGCAAGACAC	1260
1201	CAGGATTTTCATATAGTGAAGAGTGTCTGTGTGTCCTTTCATATTTGGCAAGACAC	1260
1261	AAATGAGCTAAGATTTACTGTTTTTCAGTTTCATCGATTTTCTATTTAGAAAACTGTGT	1320
1261	AAATGAGCTAAGATTTACTGTTTTTCAGTTTCATCGATTTTCTATTTAGAAAACTGTGT	1320

RESULT 3

```

1  RESULT 3
2  US-10-127-551-1
3  ; Sequence 1, Application US/10127551
4  ; Publication No. US20020120123A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Rosen et al.
7  ; TITLE OF INVENTION: Vascular Endothelial Growth Factor 2
8  ; FILE REFERENCE: PF112P1
9  ; CURRENT APPLICATION NUMBER: US/10/127,551
10 ; CURRENT FILING DATE: 2002-04-23
11 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/465,968
12 ; PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-06
13 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/207,550
14 ; PRIOR FILING DATE: EARLIER FILING DATE: 1994-03-08
15 ; NUMBER OF SEQ ID NOS: 10
16 ; SOFTWARE: PatentIn Ver. 2.0
17 ; SEQ ID NO 1
18 ; LENGTH: 1674
19 ; TYPE: DNA
20 ; ORGANISM: Homo sapiens
21 ; FEATURE:
22 ; NAME/KEY: CDS
23 ; LOCATION: (12)..(1268)
24 ; FEATURE:
25 ; NAME/KEY: sig_peptide
26 ; LOCATION: (12)..(149)
27 ; FEATURE:
28 ; NAME/KEY: mat_peptide
29 ; LOCATION: (150)..(1268)
30 ;
31 ; US-10-127-551-1

```

Query Match	100.0%	Score 1674;	DB 13;	Length 1674;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 1674;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	GTCTTCACCATGCACTCGCTGGGCTTCTTCTGTGTGGCGTGTCTCTCTCTCGCGCGTG	60
Db	1	GTCTTCACCATGCACTCGCTGGGCTTCTTCTGTGTGGCGTGTCTCTCTCTCGCGCGTG	60
Qy	61	CGCTGCTCCCGGGTCTTCGAGAGCGCCGCGCGCGCGCGCTTCGAGTCGGAATCG	120
Db	61	CGCTGCTCCCGGGTCTTCGAGAGCGCCGCGCGCGCGCGCTTCGAGTCGGAATCG	120
Qy	121	ACCTCTCGGACGGAGGCCCGACCGGGCGAGCGCATCGGCTTATGCAAGCAAGATCTCG	180
Db	121	ACCTCTCGGACGGAGGCCCGACCGGGCGAGCGCATCGGCTTATGCAAGCAAGATCTCG	180

QY 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTGACTTACCCAGAAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCATGACTGACTTACCCAGAAT 240
QY 241 ATTGGAATAATGTACAGTGTGTCAGTGTAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
Db 241 ATTGGAATAATGTACAGTGTGTCAGTGTAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
QY 301 CCAACCTCAACTCAAGGACAGAGACTATAAAATTTGTCGAGCACATTAATAATACAG 360
Db 301 CCAACCTCAACTCAAGGACAGAGACTATAAAATTTGTCGAGCACATTAATAATACAG 360
QY 361 AGATCTTGAAAGTATTGATAATCACTGGAGAAAGACTCAATGATGCCAGGGAGTGT 420
Db 361 AGATCTTGAAAGTATTGATAATCACTGGAGAAAGACTCAATGATGCCAGGGAGTGT 420
QY 421 GTATAGATGTGGGAGAGAGTTTGGAGTCGCGAACAACACCTCTTTAAACCTCCCATGTG 480
Db 421 GTATAGATGTGGGAGAGAGTTTGGAGTCGCGAACAACACCTCTTTAAACCTCCCATGTG 480
QY 481 TGTCGGTCTACAGATGTGGGGTTGCTGCAATAGTGAAGGGCTGCAATGATGAACACCA 540
Db 481 TGTCGGTCTACAGATGTGGGGTTGCTGCAATAGTGAAGGGCTGCAATGATGAACACCA 540
QY 541 GCAGAGCTACTCAGCAAGACGTTATTGAAATTACAGTGCCTCTCTCAAGGGCCCCA 600
Db 541 GCAGAGCTACTCAGCAAGACGTTATTGAAATTACAGTGCCTCTCTCAAGGGCCCCA 600
QY 601 AACAGTAACAATCAGTTTGGCAATCACACTTCTGCGGATGATGCTTAAACTGGATG 660
Db 601 AACAGTAACAATCAGTTTGGCAATCACACTTCTGCGGATGATGCTTAAACTGGATG 660
QY 661 TTTACAGACAGTTTCAATCATTATAGAGTTCCTGCGGAGCAACACTACACAGTGC 720
Db 661 TTTACAGACAGTTTCAATCATTATAGAGTTCCTGCGGAGCAACACTACACAGTGC 720
QY 721 AGGAGCGAACAAGACCTGCCCAACCAATTACATGTGGAATAATCATCTGCAGATGCC 780
Db 721 AGGAGCGAACAAGACCTGCCCAACCAATTACATGTGGAATAATCATCTGCAGATGCC 780
QY 781 TGGCTCAGGAGATTTTATGTTTCTCGATGCTGAGATGACTCAACAGATGGATTCC 840
Db 781 TGGCTCAGGAGATTTTATGTTTCTCGATGCTGAGATGACTCAACAGATGGATTCC 840
QY 841 ATGACATCTGTGGACCAACAAGAGCTGGATGAAGAGCTGTGAGTGTCTGCAGAG 900
Db 841 ATGACATCTGTGGACCAACAAGAGCTGGATGAAGAGCTGTGAGTGTCTGCAGAG 900
QY 901 CGGGCTTCGGCTTCAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
Db 901 CGGGCTTCGGCTTCAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
QY 961 GTGTCTGTAAAAACAACCTCTCCCGAGCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Db 961 GTGTCTGTAAAAACAACCTCTCCCGAGCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
QY 1021 ACATGCGCAGTGTGTATGATAAAGAACTGTCCCGAGAAATCAACCCCTTAATCTGGAA 1080
Db 1021 ACATGCGCAGTGTGTATGATAAAGAACTGTCCCGAGAAATCAACCCCTTAATCTGGAA 1080
QY 1081 AATGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGAGAAAGAACTTC 1140
Db 1081 AATGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGAGAAAGAACTTC 1140
QY 1141 ACCACCAACATGAGCTGTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAACATGAGCTGTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
QY 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGTTGTCCTTCATATTGCAAGAACCCAC 1260
Db 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGTTGTCCTTCATATTGCAAGAACCCAC 1260
QY 1261 AAATGAGCTAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGAAAAAAGTGT 1320

Db 1261 AAATGAGCTAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGAAAAAAGTGT 1320
QY 1321 TGCCACAGTGAAGACTGTCTGTGAACAGAGACACCTTGTGGTCCATCTTAACAAAGACA 1380
Db 1321 TGCCACAGTGAAGACTGTCTGTGAACAGAGACACCTTGTGGTCCATCTTAACAAAGACA 1380
QY 1381 AAAGTCTCTCTTCTCTGAAACCATGTGGATAACTTTTACAGAAATGGACTGGAGCTCATCTG 1440
Db 1381 AAAGTCTCTCTTCTCTGAAACCATGTGGATAACTTTTACAGAAATGGACTGGAGCTCATCTG 1440
QY 1441 CAAAAGGCTCTGTGAAAGACTGTTTTCTGCCAATGACCAACAGACCAAGATTTTCTC 1500
Db 1441 CAAAAGGCTCTGTGAAAGACTGTTTTCTGCCAATGACCAACAGACCAAGATTTTCTC 1500
QY 1501 TTGTGATTTCTTTAAAGAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1560
Db 1501 TTGTGATTTCTTTAAAGAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1560
QY 1561 ATTCAATTTTATAGCAACAACAATTTGTAACACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAACAATTTGTAACACTCAGTGTGATCAATATTTTATATCAT 1620
QY 1621 GCAAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1674
Db 1621 GCAAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1674

RESULT 4

US-10-084-488-1
; Sequence 1, Application US/10084488
; Publication No. US20030028007A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,488
; FILING DATE: 28-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/623,725
; FILING DATE: 07-Sep-2000
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: 13-MAR-1998
; APPLICATION NUMBER: US/09/107,997
; FILING DATE: 30-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: MICHELE M. WALES
; REGISTRATION NUMBER: 43,975
; REFERENCE/DOCKET NUMBER: PF112PCT3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301)309-8504
; TELEFAX: (301)309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1674 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

FEATURE: NAME/KEY: sig_peptide
LOCATION: 12...80
FEATURE: NAME/KEY: mat_peptide
LOCATION: 81...1268
FEATURE: NAME/KEY: CDS
LOCATION: 12...1268
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-084-488-1

Query Match 100.0%; Score 1674; DB 14; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTGGCGGTG 60
DB 1 GTCTTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTGGCGGTG 60

QY 61 CGTGTCTCCGGGTCTCTCGAGGGCCCGCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120
DB 61 CGTGTCTCCGGGTCTCTCGAGGGCCCGCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120

QY 121 ACCTCTCGGACGCGGAGCCGCGCGGCGAGGCCACGCGCTTATGCAAGCAAGATCTGG 180
DB 121 ACCTCTCGGACGCGGAGCCGCGCGGCGAGGCCACGCGCTTATGCAAGCAAGATCTGG 180

QY 181 AGGAGCAGTACGGTCTGTGTCCAGTGTAGATGAATCTATGACTGTCTTACCCAGAT 240
DB 181 AGGAGCAGTACGGTCTGTGTCCAGTGTAGATGAATCTATGACTGTCTTACCCAGAT 240

QY 241 ATTGGAATGTACAGTGTACAGTGTACAGTGTACAGTGTACAGTGTACAGTGTACAG 300
DB 241 ATTGGAATGTACAGTGTACAGTGTACAGTGTACAGTGTACAGTGTACAGTGTACAG 300

QY 301 CCAACTCAACTCAAGGACAGAGACTATATAATTTGCTGCAGCAATATATAATACAG 360
DB 301 CCAACTCAACTCAAGGACAGAGACTATATAATTTGCTGCAGCAATATATAATACAG 360

QY 361 AGATCTTGAAGTATTGATATAGTGGAGAGAGACTCAATGCGATGCCGAGGAGT 420
DB 361 AGATCTTGAAGTATTGATATAGTGGAGAGAGACTCAATGCGATGCCGAGGAGT 420

QY 421 GTATAGATGTGGGAGAGGTTGGAGTGCAGCAAAACACCTTCTTTAAACCTCCATGTG 480
DB 421 GTATAGATGTGGGAGAGGTTGGAGTGCAGCAAAACACCTTCTTTAAACCTCCATGTG 480

QY 481 TGTCCGTACAGATGTGGGAGTGTGCAATAGTGGAGAGAGTGTGCAATAGTGGAGAGT 540
DB 481 TGTCCGTACAGATGTGGGAGTGTGCAATAGTGGAGAGAGTGTGCAATAGTGGAGAGT 540

QY 541 GCACGAGTACCTCAGCAAGAGCTTATTTGAATACAGTGTCTCTCTCAAGGCCCA 600
DB 541 GCACGAGTACCTCAGCAAGAGCTTATTTGAATACAGTGTCTCTCTCAAGGCCCA 600

QY 601 AACCACTAACTACGTTTGGCAATCACTTCTGCGGATGCAATGCTAACTGATG 660
DB 601 AACCACTAACTACGTTTGGCAATCACTTCTGCGGATGCAATGCTAACTGATG 660

QY 661 TTTACAGACAGTTCATTCATATTAGAGCTTCCCTGCGGACCACTACCAAGTGTG 720
DB 661 TTTACAGACAGTTCATTCATATTAGAGCTTCCCTGCGGACCACTACCAAGTGTG 720

QY 721 AGGACGCAAGACAGCTGCCCAACCAATACATGTGGAATATACATCTCCAGATGCC 780
DB 721 AGGACGCAAGACAGCTGCCCAACCAATACATGTGGAATATACATCTCCAGATGCC 780

QY 781 TGCTCAGGAGATTTATGTTTCTCGGATGCTGAGATGACTCAAGAGTGTATCC 840
DB 781 TGCTCAGGAGATTTATGTTTCTCGGATGCTGAGATGACTCAAGAGTGTATCC 840

QY 841 ATGACATCTGTGACCAAAACAGGAGCTGGATGAAGAGACCTGTCTAGTGTCTGAGAG 900

DB 841 ATGACATCTGTGACCAAAACAGGAGCTGGATGAAGAGACCTGTCTAGTGTCTGAGAG 900
QY 901 CGGGGCTTCGGCTGCGCAGCTGTGGACCCCAACAGAACTAGACAGAAATCTATGCCAGT 960
DB 901 CGGGGCTTCGGCTGCGCAGCTGTGGACCCCAACAGAACTAGACAGAAATCTATGCCAGT 960

QY 961 GTGTCTGTAATAACAAACTCTTCCAGCCCAATGTGGGCGCAACCGAGATTTGATGAAA 1020
DB 961 GTGTCTGTAATAACAAACTCTTCCAGCCCAATGTGGGCGCAACCGAGATTTGATGAAA 1020

QY 1021 ACACATGCCAGTGTATGTATAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTGGAA 1080
DB 1021 ACACATGCCAGTGTATGTATAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTGGAA 1080

QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGGAAAGAGTTCC 1140
DB 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGGAAAGAGTTCC 1140

QY 1141 ACCACCAAAACATGAGCTGTTACAGAGCGGCATGTACGAACCCGCAAGAGCTTGTGAGC 1200
DB 1141 ACCACCAAAACATGAGCTGTTACAGAGCGGCATGTACGAACCCGCAAGAGCTTGTGAGC 1200

QY 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGTGCTTTCATATGTGCAAGAGCCAC 1260
DB 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGTGCTTTCATATGTGCAAGAGCCAC 1260

QY 1261 AATGAGCTAAGATGTGCTGTTTCCAGTTTCTGATTTCTTATGTAAGAACTGTGT 1320
DB 1261 AATGAGCTAAGATGTGCTGTTTCCAGTTTCTGATTTCTGATTTCTTATGTAAGAACTGTGT 1320

QY 1321 TGCCACAGTAAAGTGTGCTGAACAGAGAGACCTTGTGGTCCATGTCTAACAAAGACA 1380
DB 1321 TGCCACAGTAAAGTGTGCTGAACAGAGAGACCTTGTGGTCCATGTCTAACAAAGACA 1380

QY 1381 AAGTCTGCTTCTTCCAGTGTGATTAACCTTACAGAAATGCTGAGCTCATCTG 1440
DB 1381 AAGTCTGCTTCTTCCAGTGTGATTAACCTTACAGAAATGCTGAGCTCATCTG 1440

QY 1441 CAAAGGCTCTTGTAAAGACTGTTTCTGCAATGACCAACAGCAAGATTTTCTC 1500
DB 1441 CAAAGGCTCTTGTAAAGACTGTTTCTGCAATGACCAACAGCAAGATTTTCTC 1500

QY 1501 TTGTGATTTCTTAAAGAAATGATATATTAATTTTCCACTTAAATAATTTGTTTCTG 1560
DB 1501 TTGTGATTTCTTAAAGAAATGATATATTAATTTTCCACTTAAATAATTTGTTTCTG 1560

QY 1561 ATTCAATTTTATGACCAACAAATTTGTAATAATTTTATATATATATATATAT 1620
DB 1561 ATTCAATTTTATGACCAACAAATTTGTAATAATTTTATATATATATATATATAT 1620

QY 1621 GCAAAATATCTTAAATAATAAATGAAATTTGATTTTATAAAAAA 1674
DB 1621 GCAAAATATCTTAAATAATAAATGAAATTTGATTTTATAAAAAA 1674

RESULT 5
US-10-120-398-1
; Sequence 1, Application US/10120398
; Publication No. US20030170786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig, et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: Pfl12P7
; CURRENT APPLICATION NUMBER: US/10/120,398
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/283,408
; PRIOR FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: sig_peptide
LOCATION: (12)..(80)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (12)..(1268)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (81)..()
OTHER INFORMATION:
US-10-120-398-1

Query Match 100.0%; Score 1674; DB 14; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTCCAGATGCACTGCTGGGCTTCTTCTGTGGGCTTCTTCTGCTCGCGCTG 60
Db 1 GTCTTCCAGATGCACTGCTGGGCTTCTTCTGTGGGCTTCTTCTGCTCGCGCTG 60

QY 61 CGTGTCTCCCGGTCCTCGGAGCGCCCGCCGCGCTTCGAGTCGGAATCG 120
Db 61 CGTGTCTCCCGGTCCTCGGAGCGCCCGCCGCGCTTCGAGTCGGAATCG 120

QY 121 ACCTCTCGAGCGGAGCCGAGCGCGGCGAGCCACGGCTTATGCAAGAAAGATCTGG 180
Db 121 ACCTCTCGAGCGGAGCCGAGCGCGGCGAGCCACGGCTTATGCAAGAAAGATCTGG 180

QY 181 AGGAGCAGTTACGGTCTGTCTGAGTGTAGTGAATCTCATGACTGTACTTACCAGAAAT 240
Db 181 AGGAGCAGTTACGGTCTGTCTGAGTGTAGTGAATCTCATGACTGTACTTACCAGAAAT 240

QY 241 ATTGGAAAATGTACAAGTGTGAGTGAAGAGAGCTCAATGATGCTGCCAACAACAGAACAGG 300
Db 241 ATTGGAAAATGTACAAGTGTGAGTGAAGAGAGAGCTCAATGATGCTGCCAACAACAGAACAGG 300

QY 301 CCAACCTCACTCAAGCAGAGAGAGACTATAAATTTGCTGAGCAGACATTAATATACAG 360
Db 301 CCAACCTCACTCAAGCAGAGAGAGACTATAAATTTGCTGAGCAGACATTAATATACAG 360

QY 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCTGCCAAGAGGTGT 420
Db 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCTGCCAAGAGGTGT 420

QY 421 GTATAGATGGGGAAGAGTTGGAGTCCGACAAACACCTTCTTTAACTCCATGTG 480
Db 421 GTATAGATGGGGAAGAGTTGGAGTCCGACAAACACCTTCTTTAACTCCATGTG 480

QY 481 TGTCCGTCTACAGATGTGGGTTGCTGCAATAGTGGGGCTGCAGTGCATGAACACCA 540
Db 481 TGTCCGTCTACAGATGTGGGTTGCTGCAATAGTGGGGCTGCAGTGCATGAACACCA 540

QY 541 GCACGAGTACCTCAGCAAGAGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Db 541 GCACGAGTACCTCAGCAAGAGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600

QY 601 AACCAAGTAAACATCAGTTTGGCAATCACACTTCTCCCGATGCTGTCTAACTGGATG 660
Db 601 AACCAAGTAAACATCAGTTTGGCAATCACACTTCTCCCGATGCTGTCTAACTGGATG 660

QY 661 TTTACAGAGTTTCATTCATTAATTAAGAGTTTCCCTGCCAGCAACACTACACAGTGTC 720
Db 661 TTTACAGAGTTTCATTCATTAATTAAGAGTTTCCCTGCCAGCAACACTACACAGTGTC 720

QY 721 AGGAGCGGAACAGACCTGCCCAACCAATACATGCTGAATATCACTCTGAGATGCC 780
Db 721 AGGAGCGGAACAGACCTGCCCAACCAATATACATGCTGAATATCACTCTGAGATGCC 780

QY 781 TGGCTCAGGAAGATTTTATGTTTCTCCGATGCTGGAGATGACTCAACAGATGGATTCC 840
Db 781 TGGCTCAGGAAGATTTTATGTTTCTCCGATGCTGGAGATGACTCAACAGATGGATTCC 840

Db 781 TGGCTCAGGAAGATTTTATGTTTCTCCGATGCTGGAGATGACTCAACAGATGGATTCC 840
QY 841 ATGACATCTGTGGACCAAAACAAGGAGCTGGATGAAGAGACCTGTGAGTGTCTGCGAGAG 900
Db 841 ATGACATCTGTGGACCAAAACAAGGAGCTGGATGAAGAGACCTGTGAGTGTCTGCGAGAG 900
QY 901 CGGGCTTCCGCTGCGAGCTGTGAGCCCAAGAACTAGACAGAAATCTATGCGCAGT 960
Db 901 CGGGCTTCCGCTGCGAGCTGTGAGCCCAAGAACTAGACAGAAATCTATGCGCAGT 960
QY 961 GTGCTGTAAAAACAACACTCTTCCCGAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
Db 961 GTGCTGTAAAAACAACACTCTTCCCGAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
QY 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCGCCAGAAATCAACCCCTAAATCCTGGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCGCCAGAAATCAACCCCTAAATCCTGGAA 1080
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGATTTCC 1140
Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGATTTCC 1140
QY 1141 ACCACCAACATGAGCTGTGTACAGCGCCCATGTACGAAACCGCCAGAGCTTTGAGC 1200
Db 1141 ACCACCAACATGAGCTGTGTACAGCGCCCATGTACGAAACCGCCAGAGCTTTGAGC 1200
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTTCCCTTCATATTTGCAAGAGCCAC 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTTCCCTTCATATTTGCAAGAGCCAC 1260
QY 1261 AAATCAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGAAAACTGTGT 1320
Db 1261 AAATCAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGAAAACTGTGT 1320
QY 1321 TGCCACAGTAGAAGTGTGTGAACAGAGAGACCTTGTGGTCCATGCTTAACAGAGACA 1380
Db 1321 TGCCACAGTAGAAGTGTGTGAACAGAGAGACCTTGTGGGTCCATGCTTAACAGAGACA 1380
QY 1381 AAGTCTCTCTTCTGAAACCATGTGGATTAATTTACAGAAATGAGCTGGAGCTCATCTG 1440
Db 1381 AAGTCTCTCTTCTGAAACCATGTGGATTAATTTACAGAAATGAGCTGGAGCTCATCTG 1440
QY 1441 CAAAGGCTCTGTGAAGAGCTGGTTTTCTGCCAATGACCAAGACGCAAGATTTCTCTC 1500
Db 1441 CAAAGGCTCTGTGAAGAGCTGGTTTTCTGCCAATGACCAAGACGCAAGATTTCTCTC 1500
QY 1501 TTGTGATTTCTTTAAAGAGTGTATATATTTTCCACTAAAAATATTGTTTCTGC 1560
Db 1501 TTGTGATTTCTTTAAAGAGTGTATATATTTTCCACTAAAAATATTGTTTCTGC 1560
QY 1561 ATTCAATTTTATAGCAACAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
QY 1621 GCAAAATATGTTTAAATTAATTAATGAAATTTGATTTTAAAAAATTTTAAAAA 1674
Db 1621 GCAAAATATGTTTAAATTAATTAATGAAATTTGATTTTAAAAAATTTTAAAAA 1674

RESULT 6
US-10-120-414-1
; Sequence 1, Application US/10120414
; Publication No. US20030175274A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig, et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P9
; CURRENT APPLICATION NUMBER: US/10/120,414
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/283,385
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/350,366
; PRIOR FILING DATE: 2002-01-24

```

; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (12)..(60)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1269)
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: (81)..(1)
; OTHER INFORMATION:
US-10-120-414-1
Query Match 100.0%; Score 1674; DB 14; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCCTTCACCATGCACTCGCTGGCTTCTCTCTGTGGGCTGTCTCTGTCGCCGCTG 60
DB 1 GTCCTTCACCATGCACTCGCTGGCTTCTCTCTGTGGGCTGTCTCTGTCGCCGCTG 60
QY 61 CGCTGCTCCCGGTCCTCGCAGGCGCCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
DB 61 CGCTGCTCCCGGTCCTCGCAGGCGCCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
QY 121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 180
DB 121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 180
QY 181 AGGACAGTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240
DB 181 AGGACAGTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240
QY 241 ATTGAAAATGTACAAAGTGTAGCTAAGAAAGAGGCTGGCAACATACAGAAACAGG 300
DB 241 ATTGAAAATGTACAAAGTGTAGCTAAGAAAGAGGCTGGCAACATACAGAAACAGG 300
QY 301 CCAACTCACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAGCATTAATACAG 360
DB 301 CCAACTCACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAGCATTAATACAG 360
QY 361 AGATCTTGAAAATGATTAATAGTGTGAGAAAGACTCAATGTCATGCCAGGAGGTGT 420
DB 361 AGATCTTGAAAATGATTAATAGTGTGAGAAAGACTCAATGTCATGCCAGGAGGTGT 420
QY 421 GTATAGATGTGGGAGAGAGTTTGAGTTCGGACAAACACCTCTTTTAACTCCATGTG 480
DB 421 GTATAGATGTGGGAGAGAGTTTGAGTTCGGACAAACACCTCTTTTAACTCCATGTG 480
QY 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGTGAGGGGTGAGTGCATGACCA 540
DB 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGTGAGGGGTGAGTGCATGACCA 540
QY 541 GCACGAGCTACCTCAGCAAGAAGCTTTATTTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
DB 541 GCACGAGCTACCTCAGCAAGAAGCTTTATTTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
QY 601 ACCAGTACATCACTTTTGGCAATCACTTCTGCGGATGATGCTTAACTGGATG 660
DB 601 ACCAGTACATCACTTTTGGCAATCACTTCTGCGGATGATGCTTAACTGGATG 660
QY 661 TTTACAGCAAGTTTCACTTATTTAGAGTTCCTGCCAGCAACATCACACAGTGTG 720
DB 661 TTTACAGCAAGTTTCACTTATTTAGAGTTCCTGCCAGCAACATCACACAGTGTG 720
QY 721 AGGACGGACACAGACTGCCCCACCAATTAATGATGTGGAATTAATCACATCTCAGATGCC 780
```

```

DB 721 AGGACGGACACAGACTGCCCCACCAATTAATGATGTGGAATTAATCACATCTCAGATGCC 780
QY 781 TGGCTCAGGAAGATTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTCC 840
DB 781 TGGCTCAGGAAGATTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTCC 840
QY 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTGCTGCTCCAGAG 900
DB 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTGCTGCTCCAGAG 900
QY 901 CGGGGCTTCGGCTCGCAGCTGTGACCCCAAAAGAACTAGACAGAACTCATGCCAGT 960
DB 901 CGGGGCTTCGGCTCGCAGCTGTGACCCCAAAAGAACTAGACAGAACTCATGCCAGT 960
QY 961 GTGTCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
DB 961 GTGTCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
QY 1021 ACACATGCCAGTGTGTATGTAAGAACTCTGACCCCAAGAAATCAACCCCTTAATCTGGA 1080
DB 1021 ACACATGCCAGTGTGTATGTAAGAACTCTGACCCCAAGAAATCAACCCCTTAATCTGGA 1080
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGTTCC 1140
DB 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGTTCC 1140
QY 1141 ACCACCAACATGACGCTGTGTACAGAGCGCCATGTACGAAACCGCAGAGGCTGTGAGC 1200
DB 1141 ACCACCAACATGACGCTGTGTACAGAGCGCCATGTACGAAACCGCAGAGGCTGTGAGC 1200
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTCGTGTGTGCTCCCTTCATATTTGGCAAGCCAC 1260
DB 1201 CAGGATTTTCATATAGTGAAGAGTGTGTCGTGTGTGCTCCCTTCATATTTGGCAAGCCAC 1260
QY 1261 AATGAGCTAAGATTTGCTGTTTCCAGTTCATCGATTTCTTATTTATGGAATCTGTG 1320
DB 1261 AATGAGCTAAGATTTGCTGTTTCCAGTTCATCGATTTCTTATTTATGGAATCTGTG 1320
QY 1321 TGCCACAGTAGAATCTGTGTGAACAGAGAGACCTTGTGGGTCCATGTAAACAAGACA 1380
DB 1321 TGCCACAGTAGAATCTGTGTGAACAGAGAGACCTTGTGGGTCCATGTAAACAAGACA 1380
QY 1381 AAAGTCTGTCTTCCCTGAACCATGTGGAATACTTTACAGAAATGGAAGTCACTG 1440
DB 1381 AAAGTCTGTCTTCCCTGAACCATGTGGAATACTTTACAGAAATGGAAGTCACTG 1440
QY 1441 CAAAGGCTCTGTGAAGAGCTGTTTCTGCCAATGACCAACAGCCAGATTTTCTCTC 1500
DB 1441 CAAAGGCTCTGTGAAGAGCTGTTTCTGCCAATGACCAACAGCCAGATTTTCTCTC 1500
QY 1501 TTGTGATTTCTTTAAAGAACTGATATAATTTTCCACTTAAATAATTTGTTTCTGC 1560
DB 1501 TTGTGATTTCTTTAAAGAACTGATATAATTTTCCACTTAAATAATTTGTTTCTGC 1560
QY 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCACTGATCATATTTTATATCAT 1620
DB 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCACTGATCATATTTTATATCAT 1620
QY 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTTATAAAAAAATTTTAAAAA 1674
DB 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTTATAAAAAAATTTTAAAAA 1674
```

RESULT 7
US-10-120-377-1
; Sequence 1, Application US/10120377
; Publication No. US20030176674A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Craig, et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PF112P8
; CURRENT APPLICATION NUMBER: US/10/120,377

APPLICANT: Susan Stuart
APPLICANT: Jed G. Nuchtern
APPLICANT: Sharon E. Pion
APPLICANT: Jason M. Shohet
TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
FILE REFERENCE: PA-0046 US
CURRENT APPLICATION NUMBER: 2002-02-25
CURRENT FILING DATE: 2002-02-25
PRIOR APPLICATION NUMBER: 60/270,784
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 365
SOFTWARE: PERL Program
SEQ ID NO 98
LENGTH: 1786
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20030119009A1 550425CB1
US-10-084-817-98

Query Match 99.4%; Score 1664; DB 14; Length 1786;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GTCCTTCCACCATGCACTCGTGGGCTTCTCTGTGGGCTGTTCTCTGCTCGCGCTG 60
Db 115 GTCCTTCCACCATGCACTCGTGGGCTTCTCTGTGGGCTGTTCTCTGCTCGCGCTG 174

Qy 61 CGTGTCTCCGGTCTCTCGGAGCGCCCGCGCGCGCGCGCTTCGAGTCGGACTCG 120
Db 175 CGTGTCTCCGGTCTCTCGGAGCGCCCGCGCGCGCGCTTCGAGTCGGACTCG 234

Qy 121 ACCTCTCGGACGCGGAGCGCCGAGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
Db 235 ACCTCTCGGACGCGGAGCGCCGAGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 294

Qy 181 AGGAGCTAGTACGCTGTGTGTGCTGATGATGATGATGATGATGATGATGATGAT 240
Db 295 AGGAGCTAGTACGCTGTGTGTGCTGATGATGATGATGATGATGATGATGATGAT 354

Qy 241 ATTGGAATATCTACAGTGTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 355 ATTGGAATATCTACAGTGTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 414

Qy 301 CCAACCTCAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 415 CCAACCTCAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 474

Qy 361 AGATCTTGAAGATATTGATTAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 475 AGATCTTGAAGATATTGATTAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 534

Qy 421 GTATAGATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 535 GTATAGATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594

Qy 481 TGTCGCTACAGATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 595 TGTCGCTACAGATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 654

Qy 541 GCAGAGCTACCTCAGCAAGACGTTATTTGCAATTAAGGAGGAGGAGGAGGAGGAG 600
Db 655 GCAGAGCTACCTCAGCAAGACGTTATTTGCAATTAAGGAGGAGGAGGAGGAGGAG 714

Qy 601 AACAGTAACAATCAGTTTGGCAATCACAATCTCTCTGCGATGATGCTTAACTGGATG 660
Db 715 AACAGTAACAATCAGTTTGGCAATCACAATCTCTCTGCGATGATGCTTAACTGGATG 774

Qy 661 TTTACAGACAGTTTCATTCATTTAGAGGTTCCCTGCGAGCAACACTACACAGTTC 720
Db 775 TTTACAGACAGTTTCATTCATTTAGAGGTTCCCTGCGAGCAACACTACACAGTTC 834

RESULT 10
US-10-346-802-2/c
; Sequence 2, Application US/10346802
; Publication No. US20030166873A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: VEGF-RELATED PROTEIN

Qy 721 AGGAGCGAACAAGACCTGCCACCAATTAATGGAATATCAATCTCGAGATGCC 780
Db 835 AGGAGCGAACAAGACCTGCCACCAATTAATGGAATATCAATCTCGAGATGCC 894

Qy 781 TGGCTCAGGAAGATTTATGTTTCTCGGATGTGGAGATGACTCAACAGATGATGCC 840
Db 895 TGGCTCAGGAAGATTTATGTTTCTCGGATGTGGAGATGACTCAACAGATGATGCC 954

Qy 841 ATGACATCTGTGGACCAACAAGAGCTGGATGAAGACCTGTCTGAGTGTCTGACAG 900
Db 955 ATGACATCTGTGGACCAACAAGAGCTGGATGAAGACCTGTCTGAGTGTCTGACAG 1014

Qy 901 CGGGGCTTCCGCTGCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1015 CGGGGCTTCCGCTGCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 1074

Qy 961 GTGTCTGTAAAAAACHAATCTTCCCGACCAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
Db 1075 GTGTCTGTAAAAAACHAATCTTCCCGACCAATGTGGGCGCAACCGAGAAATTTGATGAAA 1134

Qy 1021 ACACATGCCAGTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGAAA 1080
Db 1135 ACACATGCCAGTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGAAA 1194

Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAAAGTTCC 1140
Db 1195 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAAAGTTCC 1254

Qy 1141 ACCCAACCATGAGCTGTGTACAGCGGCATGTACGACCGCCAGAGGCTTGTGAGC 1200
Db 1255 ACCCAACCATGAGCTGTGTACAGCGGCATGTACGACCGCCAGAGGCTTGTGAGC 1314

Qy 1201 CAGGATTTTATATAGTGAAGAGTGTGCTGTTTCCAGTTTCTATTTTGGCAAGACCCAC 1260
Db 1315 CAGGATTTTATATAGTGAAGAGTGTGCTGTTTCCAGTTTCTATTTTGGCAAGACCCAC 1374

Qy 1261 AAATGAGCTAAGATGTACTGTTTCCAGTTTCTATTTTGGAAACTGTGT 1320
Db 1375 AAATGAGCTAAGATGTACTGTTTCCAGTTTCTATTTTGGAAACTGTGT 1434

Qy 1321 TGCCACAGTGAACCTGTGTGAACAGAGAGACCTTGTGGGTGCATGCTAACAAAGACA 1380
Db 1435 TGCCACAGTGAACCTGTGTGAACAGAGAGACCTTGTGGGTGCATGCTAACAAAGACA 1494

Qy 1381 AAAGTCTCTTCTCTGAAACCATGTGGATAACTTTACAGAAATGAGCTGAGAGCTCATCTG 1440
Db 1495 AAAGTCTCTTCTCTGAAACCATGTGGATAACTTTACAGAAATGAGCTGAGAGCTCATCTG 1554

Qy 1441 CAAAGGCTCTTGTGAAGAGCTGTTTCTGCCCAATGACCAACAGACCAAGATTTCTCTC 1500
Db 1555 CAAAGGCTCTTGTGAAGAGCTGTTTCTGCCCAATGACCAACAGACCAAGATTTCTCTC 1614

Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATATTTTATTTTCCACTAAAAATATTTTCTGTC 1560
Db 1615 TTGTGATTTCTTTAAAGAAATGACTATATATTTTATTTTCCACTAAAAATATTTTCTGTC 1674

Qy 1561 ATTCAATTTTATAGCAACAACAATTTGTAAGAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1675 ATTCAATTTTATAGCAACAACAATTTGTAAGAACTCACTGTGATCAATATTTTATATCAT 1734

Qy 1621 GCATAATATGTTTAAATATAATGAAATTTGATTTTATATAAAAAA 1672
Db 1735 GCATAATATGTTTAAATATAATGAAATTTGATTTTATATAAAAAA 1786

FILE REFERENCE: P0963R1D1
CURRENT APPLICATION NUMBER: US/10/346,802
PRIOR FILING DATE: 2003-01-17
PRIOR APPLICATION NUMBER: US/09/313,299B
PRIOR FILING DATE: 1999-05-17
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 08/706,054
PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-30
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/003,491
PRIOR FILING DATE: EARLIER FILING DATE: 1995-09-08
NUMBER OF SEQ ID NOS: 12
SEQ ID NO 2
LENGTH: 2031
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: Human
LOCATION: 1-2031
OTHER INFORMATION: Sequence source: complement to SEQ ID NO. 1
US-10-346-802-2

Query Match 99.3%; Score 1663; DB 14; Length 2031;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1666; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCCTCCACCATGACCTGCGTGGGCTTCTCTGTGGCGGTCTCTGTCTGCTCGCGCTG 60
DB 1671 GTCCCTCCACCATGACCTGCTGGGCTTCTCTGTGGCGGTCTCTGTCTGCTCGCGCTG 1612

QY 61 CCGTCTCCCGGTCCTCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 1611 TCGTCTCCCGGTCCTCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1552

QY 121 ACCTCTCGAGCGCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 1551 ACCTCTCGAGCGCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1492

QY 181 AGGAGCAGTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTTACCAGAAAT 240
DB 1491 AGGAGCAGTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTTACCAGAAAT 1432

QY 241 ATTGAAATATGTACAAATGTACAGTGTACAGTGTAGTAAAGAGGCGTGGCAATACAGAGACAGG 300
DB 1431 ATTGAAATATGTACAAATGTACAGTGTACAGTGTAGTAAAGAGGCGTGGCAATACAGAGACAGG 1372

QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAGCAATATATATACAG 360
DB 1371 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAGCAATATATATACAG 1312

QY 361 AGATCTTGAAAGTATTGATATAGTGTGAGAGAAAGACTCAATGTGATGCCAGGAGGTGT 420
DB 1311 AGATCTTGAAAGTATTGATATAGTGTGAGAGAAAGACTCAATGTGATGCCAGGAGGTGT 1252

QY 421 GTATAGATGTGGGAGGAGTTGAGTTCGGAGCAACACCTTCTTTAAACCTCCATGTG 480
DB 1251 GTATAGATGTGGGAGGAGTTGAGTTCGGAGCAACACCTTCTTTAAACCTCCATGTG 1192

QY 481 TGTCCGTCTACAGATGTGGGGTTCCTGCAATAGTGTAGGCGGTGAGTGTGATGACACCA 540
DB 1191 TGTCCGTCTACAGATGTGGGGTTCCTGCAATAGTGTAGGCGGTGAGTGTGATGACACCA 1132

QY 541 GCACGAGTCTACAGAGAGGTTATTGAAATTTACAGTGGCTCTCTCAAGGCCCA 600
DB 1131 GCACGAGTCTACAGAGAGGTTATTGAAATTTACAGTGGCTCTCTCTCAAGGCCCA 1072

QY 601 AACAGTAAATCAATGATTGTCATACACTTCTGCGGATGATGCTTAACTGGATG 660
DB 1071 AACAGTAAATCAATGATTGTCATACACTTCTGCGGATGATGCTTAACTGGATG 1012

QY 661 TTATACAGCAAGTTCATTCATTTATAGAGTTCCTGCGAGCAACACTACACAGTGTG 720
DB 1011 TTATACAGCAAGTTCATTCATTTATAGAGTTCCTGCGAGCAACACTACACAGTGTG 952

QY 721 AGGAGCGAACAAGACCTGCGCCCAATTTACATGTGGAATATACATCTGCAGATGCC 780

DB 951 AGGAGCGAACAAGACCTGCCCAATTTACATGTGGAATATCACTATCGAGATGCC 892
QY 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTCGAGATGACTCAACAGATGGATTC 840
DB 891 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTCGAGATGACTCAACAGATGGATTC 832
QY 841 ATGACATCTGTGGACCAACAAAGAGCTGTGATGAAGAGACCTGTCTAGTGTGTGAGAG 900
DB 831 ATGACATCTGTGGACCAACAAAGAGCTGTGATGAAGAGACCTGTCTAGTGTGTGAGAG 772
QY 901 CGGGGCTTGGGCTGCGAGTGTGACCCCAAGAACTAGACAGAACTAGACAGAACTCATGCCAGT 960
DB 771 CGGGGCTTGGGCTGCGAGTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 712
QY 961 GTCTCTGTAAAAACAAACCTTCTCCCGAGCAATGTGGGCGCAACCCGAGAAATTTGATGAAA 1020
DB 711 GTCTCTGTAAAAACAAACCTTCTCCCGAGCAATGTGGGCGCAACCCGAGAAATTTGATGAAA 652
QY 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCGGCGCAACCCGAGAAATCAACCCCTAAATCTGGAA 1080
DB 651 ACACATGCCAGTGTGTATGTAAAGAACTGCGGCGCAACCCGAGAAATCAACCCCTAAATCTGGAA 592
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTC 1140
DB 591 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTC 532
QY 1141 ACCACAAACATGCGAGTGTACAGAGGCGCATGTACGAACCCGCAAGAGGCTTGTGAGC 1200
DB 531 ACCACAAACATGCGAGTGTACAGAGGCGCATGTACGAACCCGCAAGAGGCTTGTGAGC 472
QY 1201 CAGGATTTTCATAGTGAAGAGTGTGCTGCTGCTCCCTCATATTTGCAAGAACAC 1260
DB 471 CAGGATTTTCATAGTGAAGAGTGTGCTGCTGCTCCCTCATATTTGCAAGAACAC 412
QY 1261 AAATGAGCTAAGATTTACTGTTTCCAGTTCATCGATTTCTATTATGAAAACTGTGT 1320
DB 411 AAATGAGCTAAGATTTACTGTTTCCAGTTCATCGATTTCTATTATGAAAACTGTGT 352
QY 1321 TGCCACAGTAGACTCTCTGTGACAGAGAGACCTTGTGGTCCATGCTTACAAAGACA 1380
DB 351 TGCCACAGTAGACTCTCTGTGACAGAGAGACCTTGTGGTCCATGCTTACAAAGACA 292
QY 1381 AAAGTCTGTCTTCTTCCGAACTGTGGAATTAACCTTTACAGAAATGGAAGTGTGAGTGTG 1440
DB 291 AAAGTCTGTCTTCTTCCGAACTGTGGAATTAACCTTTACAGAAATGGAAGTGTGAGTGTG 232
QY 1441 CAAAAGGCTCTGTGAAAGCTGGTTTCTGCAATGACCAACAGCCAGAGATTTTCCTC 1500
DB 231 CAAAAGGCTCTGTGAAAGCTGGTTTCTGCAATGACCAACAGCCAGAGATTTTCCTC 172
QY 1501 TTGTGATTTCTTTAAAGAACTGACTATAATTTATTTCACCTTAAATAATTTGTTCTG 1560
DB 171 TTGTGATTTCTTTAAAGAACTGACTATAATTTATTTCACCTTAAATAATTTGTTCTG 112
QY 1561 ATTCAATTTTATAGCAACCAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT 1620
DB 111 ATTCAATTTTATAGCAACCAATTTGGTAAAACTCACTGTGATCAATATTTTATATCAT 52
QY 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTTATATAAAAAA 1671
DB 51 GCAAAATATGTTTAAATAAATGAAATTTGATTTTAAATAAAAAA 1

RESULT 11
US-09-795-006A-21
; Sequence 21, Application US/09795006A
; Patent No. US20020151680A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: MATERIALS AND METHODS INVOLVING HYBRID VASCULAR
; TITLE OF INVENTION: ENDOTHELIAL GROWTH FACTOR DNAs AND PROTEINS
; FILE REFERENCE: 28967/35977B

```

; CURRENT APPLICATION NUMBER: US/09/795,006A
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/205,331
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/185,205
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)..(1608)
US-09-795-006A-21

```

Query Match	98.7%;	Score 1651.8;	DB 9;	Length 1997;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1653; Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
QY	1	GTCTTCCACATGCACTCGCTGGCTCTCTCTGTGGGTGTTCTCTCTGCTCGCGCTG 60		
Db	341	GTCTTCCACATGCACTCGCTGGCTCTCTCTGTGGGTGTTCTCTCTGCTCGCGCTG 400		
QY	61	CGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCGCGCGCGCGCGCTG 120		
Db	401	CGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCGCGCGCGCGCGCGCTG 460		
QY	121	ACCTCTCGAGCGGAGCCCGAGCGGGCGAGCGGCACGGCTTATGCAAGCAAGATCTGG 180		
Db	461	ACCTCTCGAGCGGAGCCCGAGCGGGCGAGCGGCACGGCTTATGCAAGCAAGATCTGG 520		
QY	181	AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240		
Db	521	AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 580		
QY	241	ATTGGAAAAATGTAACAAGTGTCAAGTAAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300		
Db	581	ATTGGAAAAATGTAACAAGTGTCAAGTAAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 640		
QY	301	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGGACATATATAATACAG 360		
Db	641	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGGACATATATAATACAG 700		
QY	361	AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGCATGCCACGGGAGGTG 420		
Db	701	AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGCATGCCACGGGAGGTG 760		
QY	421	GTATAGATGTGGGAAGGAGTTTGGAGTCGGGCAAAACACCTTCTTTAAACCTCCATGTG 480		
Db	761	GTATAGATGTGGGAAGGAGTTTGGAGTCGGGCAAAACACCTTCTTTAAACCTCCATGTG 820		
QY	481	TGTCCTGTCTACAGATGTGGGGTTGCTGCAATAGTGAAGGGCTGCAGTGCATGAACACCA 540		
Db	821	TGTCCTGTCTACAGATGTGGGGTTGCTGCAATAGTGAAGGGCTGCAGTGCATGAACACCA 880		
QY	541	GCACGAGCTACCTCAGCAAGACGTTATTTGAAATATACAGTGCCTCTCTCTCAAGGCCCA 600		
Db	881	GCACGAGCTACCTCAGCAAGACGTTATTTGAAATATACAGTGCCTCTCTCTCAAGGCCCA 940		
QY	601	AACGAGTAAACAATCAGTTTGGCAATCACACTTCTCCGATGCGATGCTCTAAACTGGATG 660		
Db	941	AACGAGTAAACAATCAGTTTGGCAATCACACTTCTCCGATGCGATGCTCTCTAAACTGGATG 1000		
QY	661	TTTACAGACAAGTTTCATTCCATTATTAGACGTTTCCCTGCCAGGAACTACTACCACAGTGC 720		
Db	1001	TTTACAGACAAGTTTCATTCCATTATTAGACGTTTCCCTGCCAGGAACTACTACCACAGTGC 1060		
QY	721	AGGAGCGAACAAGCTGCCCCCACTAAATTAATGTGGAATAATCAATCTGCAGATGCC 780		
Db	1061	AGGAGCGAACAAGCTGCCCCCACTAAATTAATGTGGAATAATCAATCTGCAGATGCC 1120		

```

RESULT 12
US-09-375
; Sequenc
; Publica
; GENERAL
; APPLIC
; APPLIC
; APPLIC
; APPLIC
; TITLE
; TITLE
; FILE R

```

Qy	781	TGCGTCAGGAAGATTTTATGTTTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC	840
Db	1121	TGCGTCAGGAAGATTTTATGTTTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC	1180
Qy	841	ATGACATCTCTGGACCAAAACAAGAGCTGGATCAAGAGACCTGTCACTGTGTCTGCAGAG	900
Db	1181	ATGACATCTCTGGACCAAAACAAGAGCTGGATCAAGAGACCTGTCACTGTGTCTGCAGAG	1240
Qy	901	CGGGGCTTCGGCCTGCGAGCTGTGGACCCCAAAAGAACTTAGACAGAAACTCATATGCCAGT	960
Db	1241	CGGGGCTTCGGCCTGCGAGCTGTGGACCCCAAAAGAACTTAGACAGAAACTCATATGCCAGT	1300
Qy	961	GTGCTGTAAAAACAACCTCTCCACAGCAATGTGGGGCCACCCGAGATTTTGATGAAA	1020
Db	1301	GTGCTGTAAAAACAACCTCTCCACAGCAATGTGGGGCCACCCGAGATTTTGATGAAA	1360
Qy	1021	ACACATGCCAGTGTATGTATATAAGAAACCTGCCGCCAGAAATCAACCCCTAAATCTCGAA	1080
Db	1361	ACACATGCCAGTGTATGTATATAAGAAACCTGCCGCCAGAAATCAACCCCTAAATCTCGAA	1420
Qy	1081	AATGTGCGCTGTGAATGTACAGAAAGTCCACAGAAATGCCTTGTTTAAAGGNAAGAAGTTC	1140
Db	1421	AATGTGCGCTGTGAATGTACAGAAAGTCCACAGAAATGCCTTGTTTAAAGGNAAGAAGTTC	1480
Qy	1141	ACCACAAACATGCAGCTGTTACAGACGGCCATGTACGAACCCCGACAGAGCGCTTGTGAGC	1200
Db	1481	ACCACAAACATGCAGCTGTTACAGACGGCCATGTACGAACCCCGACAGAGCGCTTGTGAGC	1540
Qy	1201	CAGGATTTTCATATAGTAAGAAAGTGTGTGCTTGTGTGCCCTTCATATGGCAAGACCCAC	1260
Db	1541	CAGGATTTTCATATAGTAAGAAAGTGTGTGCTTGTGTGCCCTTCATATGGCAAGACCCAC	1320
Qy	1261	AAATGAGCTAAGATGTACTGTGTTTTCCAGTTTCATCATATTTCTATTATGGAAACCTGTGT	1320
Db	1601	AAATGAGCTAAGATGTACTGTGTTTTCCAGTTTCATCATATTTCTATTATGGAAACCTGTGT	1380
Qy	1321	TGCCACAGTAGAACTGTCTGTGAACAGAGAGACCCCTGTGGTGCCATGTCTAACAAAGACA	1380
Db	1661	TGCCACAGTAGAACTGTCTGTGAACAGAGAGACCCCTGTGGTGCCATGTCTAACAAAGACA	1440
Qy	1381	AAAGTCTGCTTCTTCGACCAATGTCGATTAACCTTACGANAATGGAGCTCGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTCTTCGACCAATGTCGATTAACCTTACGANAATGGAGCTCGAGCTCATCTG	1500
Qy	1441	CAAAAGGCCCTTTGTAAAAGACTGGTTTTCTGCCAAATGACCAACAGCCCAAGATTTTCCCTC	1500
Db	1781	CAAAAGGCCCTTTGTAAAAGACTGGTTTTCTGCCAAATGACCAACAGCCCAAGATTTTCCCTC	1560
Qy	1501	TTGTGATTTCTTTAAAAGAAATGACTATATAATTTATTTCCACTAAAATAATTTGTTTCTGC	1560
Db	1841	TTGTGATTTCTTTAAAAGAAATGACTATATAATTTATTTCCACTAAAATAATTTGTTTCTGC	1620
Qy	1561	ATTTCATTTTTATAGCAACCAATTCGGTAAAACTCACTGTGATCAATATTTTTATATCAT	1620
Db	1901	ATTTCATTTTTATAGCAACCAATTCGGTAAAACTCACTGTGATCAATATTTTTATATCAT	1680
Qy	1621	GC AAAATATGTTTAAAATAAAATGAAAATGTATTT 1655	
Db	1961	GC AAAATATGTTTAAAATAAAATGAAAATGTATTT 1995	

RESULT 12
US-09-375-248-3
Sequence 3, Application US/09375248
Publication No. US20030026759A1
GENERAL INFORMATION:
APPLICANT: Ferrell, Robert E.
APPLICANT: Alitalo, Kari
APPLICANT: Finogold, David N.
APPLICANT: Karhainen, Marika
TITLE OF INVENTION: SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING
TITLE OF INVENTION: THE FLT4 RECEPTOR TYROSINE KINASE (VEGFR-3)
FILE REFERENCE: 28967/35255A

CURRENT APPLICATION NUMBER: US/09/375,248
CURRENT FILING DATE: 1999-08-16
EARLIER APPLICATION NUMBER: PCT/US99/06133
EARLIER FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1997
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (352)...(1611)
FEATURE:
OTHER INFORMATION: Human cDNA for prepro-VEGF-C
US-09-375-248-3

Query Match 98.7%; Score 1651.8; DB 10; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCTTCCACCAATGCTCGTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTGGCGCTG 60
Db GTCTTCCACCAATGCTCGTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTGGCGCTG 400

QY 61 CGTGTCTCCGGGTCTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db CGTGTCTCCGGGTCTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 460

QY 121 ACCTCTCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db ACCTCTCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 520

QY 181 AGGAGCAGTACGCTGTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db AGGAGCAGTACGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 580

QY 241 ATTGGAATAATGACAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db ATTGGAATAATGACAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 640

QY 301 CCAACTCACTCAAGGACAGAGAGACTAATAAATTTGTGTGTGTGTGTGTGTGTGTGT 360
Db CCAACTCACTCAAGGACAGAGAGACTAATAAATTTGTGTGTGTGTGTGTGTGTGTGT 700

QY 361 AGATCTTGAAGATTTGATATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db AGATCTTGAAGATTTGATATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 760

QY 421 GTATAGATGTGGGAGGAGTTTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db GTATAGATGTGGGAGGAGTTTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 820

QY 481 TGTCGTCTACAGATGTGGGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db TGTCGTCTACAGATGTGGGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 880

QY 541 GCACGAGTACCTCAGCAGAGAGTTTGAATTTACAGTGTGTGTGTGTGTGTGTGTGT 600
Db GCACGAGTACCTCAGCAGAGAGTTTGAATTTACAGTGTGTGTGTGTGTGTGTGTGT 940

QY 601 AACCAATCAATAGTTTGGCAATCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db AACCAATCAATAGTTTGGCAATCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1000

QY 661 TTTACAGACAGTTTCAATTTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db TTTACAGACAGTTTCAATTTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1060

QY 721 AGGAGCGAACAGACCTGCCCAATTTACATGTGGAATTAATCACAATCTGTGTGTGT 780
Db AGGAGCGAACAGACCTGCCCAATTTACATGTGGAATTAATCACAATCTGTGTGTGTGT 1120

QY 781 TGGCTCAGGAAGATTTTATGTCTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 840
Db TGGCTCAGGAAGATTTTATGTCTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 1180

QY 841 ATGACATCTGTGGACCAACCAAGAGCTGTGATGAGAGACCTGTCACTGTGTCTGAGAG 900
Db ATGACATCTGTGGACCAACCAAGAGCTGTGATGAGAGACCTGTCACTGTGTCTGAGAG 1240

QY 901 CGGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCC 960
Db CGGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCCGGCTTCC 1300

QY 961 GTGTCTGTAAAAACAACAACTCTTCCAGCCCAATGTGGGGCCCAACCGAGAAATTTGATGAAA 1020
Db GTGTCTGTAAAAACAACAACTCTTCCAGCCCAATGTGGGGCCCAACCGAGAAATTTGATGAAA 1360

QY 1021 ACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGCCAGAAATCAACCCCTAAATCTCTGGAA 1080
Db ACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGCCAGAAATCAACCCCTAAATCTCTGGAA 1420

QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGTCTTGTAAAAAGGAAAGAGTTCC 1140
Db AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGTCTTGTAAAAAGGAAAGAGTTCC 1480

QY 1141 ACCACCAACATCATGCTGTGTACAGCGGCGCATGTAGACCGCGCATGTAGACCGCGCTGTGAGC 1200
Db ACCACCAACATCATGCTGTGTACAGCGGCGCATGTAGACCGCGCATGTAGACCGCGCTGTGAGC 1540

QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1600

QY 1261 AATGAGCTAAGATTTGACTGT 1320
Db AATGAGCTAAGATTTGACTGT 1660

QY 1321 TGCCACAGTGAAGTGTGTGTGAACAGAGAGACCTTGTGGTCCCATGTCTAACAAAGACA 1380
Db TGCCACAGTGAAGTGTGTGTGAACAGAGAGACCTTGTGGTCCCATGTCTAACAAAGACA 1720

QY 1381 AAGTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Db AAGTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1780

QY 1441 CAAAGGCTCTGTGTGAAGACTGT 1500
Db CAAAGGCTCTGTGTGAAGACTGT 1840

QY 1501 TTGTGATTTCTTAAAAAGAAATGACTATATAATTTATTTCCACTAAAAAATATTTGTCTGC 1560
Db TTGTGATTTCTTAAAAAGAAATGACTATATAATTTATTTCCACTAAAAAATATTTGTCTGC 1900

QY 1561 ATTCAATTTTATAGCAACACATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTG 1620
Db ATTCAATTTTATAGCAACACATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTG 1960

QY 1621 GCAAAATATGTTTAAAAATAAAATGAAATTTGTAATTTGTAATTTGTAATTTGTAATTTG 1655
Db GCAAAATATGTTTAAAAATAAAATGAAATTTGTAATTTGTAATTTGTAATTTGTAATTTG 1995

RESULT 13
US-10-201-386-7
Sequence 7, Application US/10201386
Publication No. US20030091567A1
GENERAL INFORMATION:
APPLICANT: Alitalia, Kari
APPLICANT: Goukov, Vladimir
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR C (VEGF-C) PROTEIN
TITLE OF INVENTION: AND GENE, MUTANTS THEREOF, AND USES THEREOF
FILE REFERENCE: 28967/34140A
CURRENT APPLICATION NUMBER: US/10/201,386
CURRENT FILING DATE: 2002-07-23

Db 1961 GCAAAATATGTTTAAATAAATGAAATGTATT 1995

RESULT 14
US-10-262-538-23
; Sequence 23, Application US/10262538
; Publication No. US20030113324A1
; GENERAL INFORMATION:
; APPLICANT: Alitalia et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23:
; LENGTH: 1997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)..(1611)
US-10-262-538-23

Query Match 98.7%; Score 1651.8; DB 14; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCTTCCACATGACACTGCTGGGCTTCTCTGTGGCGTGTCTCTGTCTCCGCGTG 60
Db 341 GTCTTCCACATGACACTGCTGGGCTTCTCTGTGGCGTGTCTCTGTCTCCGCGTG 400

QY 61 CGTCTCTCCGCGTCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 401 CGTCTCTCCGCGTCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 460

QY 121 ACCTCTCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 461 ACCTCTCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 520

QY 181 AGGAGCAGTACGCTGTGTGTCAGTGTAGATGAATCACTGATCTTACCCAGAT 240
Db 521 AGGAGCAGTACGCTGTGTGTCAGTGTAGATGAATCACTGATCTTACCCAGAT 580

QY 241 ATTGAAAATGTACAAGTGTGAGTAAAGAAAGAGGCTGCGCAACATAACAGAGAA 300
Db 581 ATTGAAAATGTACAAGTGTGAGTAAAGAAAGAGGCTGCGCAACATAACAGAGAA 640*

QY 301 CCACTCTCACTCAAGGACAGAGAGACTATAAATTTGTCGAGCACTTATAATACAG 360
Db 641 CCACTCTCACTCAAGGACAGAGAGACTATAAATTTGTCGAGCACTTATAATACAG 700

QY 361 AGATCTTCAAAAGTATTGATTAATGAGTGGAGAAAGACTCAATGATGCGGAGGAGT 420
Db 701 AGATCTTCAAAAGTATTGATTAATGAGTGGAGAAAGACTCAATGATGCGGAGGAGT 760

QY 421 GTATAGATGTGGGAGGAGTTGGAGTCGCGACAAACACCTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGTGGGAGGAGTTGGAGTCGCGACAAACACCTCTTTAAACCTCCATGTG 820

QY 481 TGTCGCTCTACAGATGTGGGAGTGTCTGCAATAGTAGGGGCTGAGTGCATGACACCA 540
Db 821 TGTCGCTCTACAGATGTGGGAGTGTCTGCAATAGTAGGGGCTGAGTGCATGACACCA 880

QY 541 GCACGAGCTACTCAGCAAGAGCTTATTGAAATTTACAGTCCCTCTCTCTCAAGGCCCA 600
Db 881 GCACGAGCTACTCAGCAAGAGCTTATTGAAATTTACAGTCCCTCTCTCTCAAGGCCCA 940

QY 601 AACCGTAAACATCATGTTTCCCAATCACACTTCTGCGGATGATGTCTAACTGATG 660
Db 941 AACCGTAAACATCATGTTTCCCAATCACACTTCTGCGGATGATGTCTAACTGATG 1000

QY 661 TTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCGGACCAACTACCAAGTGTC 720

RESULT 15
US-10-081-126-3
; Sequence 3, Application US/10081126
; Publication No. US20030180294A1
; GENERAL INFORMATION:

Db 1001 TTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCGACCAACTACCAAGTGTC 1060
QY 721 AGGAGCGGAACAAGACTGCTGCCACCAATTTACATGTGGAAATATACATCTGCAGATGCC 780
Db 1061 AGGAGCGGAACAAGACTGCTGCCACCAATTTACATGTGGAAATATACATCTGCAGATGCC 1120
QY 781 TGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATGCC 840
Db 1121 TGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATGCC 1180
QY 841 ATGACATCTGTGGACCAACAAGAGCTGATGAAGAGACCTGTGCTAGTGTCTGCAGAG 900
Db 1181 ATGACATCTGTGGACCAACAAGAGCTGATGAAGAGACCTGTGCTAGTGTCTGCAGAG 1240
QY 901 CGGGGCTTCGGCTCGCAGCTGTGGACCCCAACAAGAACTAGACAGAAACTCATGCCAGT 960
Db 1241 CGGGGCTTCGGCTCGCAGCTGTGGACCCCAACAAGAACTAGACAGAAACTCATGCCAGT 1300
QY 961 GTGCTGTAAACAACTCTTCCCGCAGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1020
Db 1301 GTGCTGTAAACAACTCTTCCCGCAGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1360
QY 1021 ACATATGCCAGTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATTCCTGGAA 1080
Db 1361 ACATATGCCAGTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATTCCTGGAA 1420
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTCC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTCC 1480
QY 1141 ACCACCAACATGTCAGCTGTTACAGACGCGCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
Db 1481 ACCACCAACATGTCAGCTGTTACAGACGCGCATGTACGAAACCGCCAGAGGCTTGTGAGC 1540
QY 1201 CAGGATTTTATATAGTGAAGAGTGTGCTGTGTGTCCTTTTATATGCGAAAGACCA 1260
Db 1541 CAGGATTTTATATAGTGAAGAGTGTGCTGTGTGTCCTTTTATATGCGAAAGACCA 1600
QY 1261 AATGAGCTAAGATGTACTGTTTCCAGTTCATCGATTTTCTTATATGGAAGAACTGTGT 1320
Db 1601 AATGAGCTAAGATGTACTGTTTCCAGTTCATCGATTTTCTTATATGGAAGAACTGTGT 1660
QY 1321 TGCCACAGTAGAACTCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGCTTAAACAAGACA 1380
Db 1661 TGCCACAGTAGAACTCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGCTTAAACAAGACA 1720
QY 1381 AAGTCTGTCTTCTGGAACCATGTGGATTAATTACAGAAATGAGTGGAGCTCATCTG 1440
Db 1721 AAGTCTGTCTTCTGGAACCATGTGGATTAATTACAGAAATGAGTGGAGCTCATCTG 1780
QY 1441 CAAAAGGCTCTTGTAAAGACTGGTTTCTGCAATGACCAACAGCCCAAGATTTTCTCTC 1500
Db 1781 CAAAAGGCTCTTGTAAAGACTGGTTTCTGCAATGACCAACAGCCCAAGATTTTCTCTC 1840
QY 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTTCCACTTAAATAATTTGTTCTG 1560
Db 1841 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTTCCACTTAAATAATTTGTTCTG 1900
QY 1561 ATTCAATTTATAGCAACAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTATAGCAACAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1660
QY 1621 GCAAAATATGTTTAAATAAATAAATGAAATTTGATT 1655
Db 1961 GCAAAATATGTTTAAATAAATAAATGAAATTTGATT 1995

APPLICANT: De Vries, Gerald W.
TITLE OF INVENTION: Methods of Extending Corneal Graft
FILE REFERENCE: P-AR 4951
CURRENT APPLICATION NUMBER: US/10/081.126
CURRENT FILING DATE: 2002-02-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2015
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (372)...(1628)
US-10-081-126-3

Query Match 98.7%; Score 1651.8; DB 14; Length 2015;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GTCCTTCACCATCACTCGCTGGGCTTCTCTGTGGGTCTCTCTCTCGCTCGCGCTG 60
Db 361 GTCCTTCACCATCACTCGCTGGGCTTCTCTGTGGGTCTCTCTCTCGCTCGCGCTG 420
61 CGTGTCTCCGGTCTCTCGGAGGCGCCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120
Db 421 CGTGTCTCCGGTCTCTCGGAGGCGCCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 480
121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 180
Db 481 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 540
181 AGGAGCAGTTACGCTCTGTGTCTAGTGAATCATGACTGTACTCTACCCAGAAT 240
Db 541 AGGAGCAGTTACGCTCTGTGTCTAGTGAATCATGACTGTACTCTACCCAGAAT 600
241 ATTGGAAATGTACAGTGTCTAGTGAATCATGACTGTACTCTACCCAGAAT 300
Db 601 ATTGGAAATGTACAGTGTCTAGTGAATCATGACTGTACTCTACCCAGAAT 360
301 CCAACCTCACTCAAGCAGAGAGAGACTATAAATTTGTCTGAGCAGACATTAATACAG 360
Db 661 CCAACCTCACTCAAGCAGAGAGAGACTATAAATTTGTCTGAGCAGACATTAATACAG 720
361 AGATCTTTGAAAGTATTGATAATCAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 420
Db 721 AGATCTTTGAAAGTATTGATAATCAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 780
421 GTATAGATGGGGAAGGTTTGGAGTCCGACAAACACTTCTTTAAACCTCCATGTG 480
Db 781 GTATAGATGGGGAAGGTTTGGAGTCCGACAAACACTTCTTTAAACCTCCATGTG 840
481 TGTCGCTCTACAGATGTGGGGTGTCTGCAATAGTGGAGGCTGCGAGTGCATGAACCA 540
Db 841 TGTCGCTCTACAGATGTGGGGTGTCTGCAATAGTGGAGGCTGCGAGTGCATGAACCA 900
541 GCAGAGCTACCTCAGCAAGACGTTATTGAAATTCAGAGCTCTCTCTCAAGGCCCA 600
Db 901 GCAGAGCTACCTCAGCAAGACGTTATTGAAATTCAGAGCTCTCTCTCAAGGCCCA 960
601 AACCAAGTAAACATCAGTTTGGCAATCACAATCTCTGCGGATCGATGCTTAACTGGATG 660
Db 961 AACCAAGTAAACATCAGTTTGGCAATCACAATCTCTGCGGATCGATGCTTAACTGGATG 1020
661 TTACAGACAGTTTCAATTCATTAATAGAGTTTCCCTGCGAGCAACACTACCAAGTGTG 720
Db 1021 TTACAGACAGTTTCAATTCATTAATAGAGTTTCCCTGCGAGCAACACTACCAAGTGTG 1080
721 AGCAGCGAACAAGACCTGCGGCAATTCAGATGGGATTAATCATCTGCGAGTGC 780
Db 1081 AGCAGCGAACAAGACCTGCGGCAATTCAGATGGGATTAATCATCTGCGAGTGC 1140

QY 781 TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
Db 1141 TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 1200
QY 841 ATGACATCTGTGGACCAACAAAGGAGCTGGATGAAGAGACCTGTCAAGTGTGTCAGAG 900
Db 1201 ATGACATCTGTGGACCAACAAAGGAGCTGGATGAAGAGACCTGTCAAGTGTGTCAGAG 1260
QY 901 CGGGCTTGGGCTCGGCTCGGAGTGTGGACCCCAAGAACTAGACAGAACTCATCCAGT 960
Db 1261 CGGGCTTGGGCTCGGCTCGGAGTGTGGACCCCAAGAACTAGACAGAACTCATCCAGT 1320
QY 961 GTGCTGTGTAACCAAACTCTTCCCGAGCCTATGTGGGCGCAACCGAGAACTTGATGAA 1020
Db 1321 GTGCTGTGTAACCAAACTCTTCCCGAGCCTATGTGGGCGCAACCGAGAACTTGATGAA 1380
QY 1021 ACACATGCGAGTGTGTATGTAAGAACTCTGCCCGAGAAATCAACCCCTAAATCCTGAA 1080
Db 1381 ACACATGCGAGTGTGTATGTAAGAACTCTGCCCGAGAAATCAACCCCTAAATCCTGAA 1440
QY 1081 ATGTGCTGTGAATGTACAGAACTGTACAGAACTGTGTAAGAACTGTGTAAGAACTGTG 1140
Db 1441 ATGTGCTGTGAATGTACAGAACTGTGTAAGAACTGTGTAAGAACTGTGTAAGAACTGTG 1500
QY 1141 ACCACCAACATGCGAGTGTGTACAGAGCGCCATGTACAGAGCGCTGTGTGAGC 1200
Db 1501 ACCACCAACATGCGAGTGTGTACAGAGCGCCATGTACAGAGCGCTGTGTGAGC 1560
QY 1201 CAGGATTTTATATAGTGAAGAGTGTGCTGTGTCCTTCTATATTTGGCAAGACAC 1260
Db 1561 CAGGATTTTATATAGTGAAGAGTGTGCTGTGTCCTTCTATATTTGGCAAGACAC 1620
QY 1261 AAATGAGCTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTTATGAAAATCTGTG 1320
Db 1621 AAATGAGCTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTTATGAAAATCTGTG 1680
QY 1321 TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCTTTTGGGTCCATGCTAACAAAGACA 1380
Db 1681 TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCTTTTGGGTCCATGCTAACAAAGACA 1740
QY 1381 AAAGTCTGTCTTCTGAGCCATGTTGATTAATTTTACAGAACTGCTGGAGCTCATCTG 1440
Db 1741 AAAGTCTGTCTTCTGAGCCATGTTGATTAATTTTACAGAACTGCTGGAGCTCATCTG 1800
QY 1441 CAAAGGCTCTTGTGAAGACTGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1500
Db 1801 CAAAGGCTCTTGTGAAGACTGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1860
QY 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTTTCACCTAAATAATTTGTTCTGC 1560
Db 1861 TTGTGATTTCTTTAAAGAGTACTATATAATTTTTCACCTAAATAATTTGTTCTGC 1920
QY 1561 ATTCAATTTTATAGCAACAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1921 ATTCAATTTTATAGCAACAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT 1980
QY 1621 GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1655
Db 1981 GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2015

Search completed: February 26, 2004, 01:43:17
Job time : 627 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 16:56:30 ; Search time 3995 Seconds
(without alignments)
12512.964 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gctcttcacatgcactcg.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_fod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1207.6	72.1	2450	11 AK047844	AK047844 Mus muscu
2	839.4	50.1	998	12 BG677890	BG677890 60262584
3	782.4	46.7	831	12 BG680919	BG680919 602628716
4	780.8	46.6	858	14 CA488579	CA488579 AGENCOURT

C	5	730.6	43.6	750	14	CA421508
C	6	729.6	43.6	748	13	BU615444
C	7	702	41.9	808	10	AW195617
C	8	698.6	41.7	717	9	AA923193
C	9	667.2	39.9	728	9	AW070215
C	10	656.2	39.2	662	12	BQ005831
C	11	649	38.8	657	13	BU619991
C	12	616	36.8	637	10	BE218812
C	13	614.6	36.7	622	14	CB854820
C	14	603	36.0	916	13	BUS10633
C	15	576	34.4	584	12	BM127186
C	16	575.4	34.4	618	9	AA151613
C	17	570	34.1	578	12	BM054372
C	18	561.2	33.5	575	10	BF669473
C	19	549.8	32.8	883	10	BF669473
C	20	527.8	31.5	564	10	BE670542
C	21	523.4	31.3	525	12	BM055226
C	22	523	31.2	535	13	BM113454
C	23	522.8	31.2	537	9	AI827929
C	24	517.2	30.9	826	10	BE58208
C	25	515.8	30.8	519	12	BM127468
C	26	514.8	30.8	570	9	AI580094
C	27	508.2	30.4	535	9	AI459340
C	28	507.4	30.3	509	9	AA425486
C	29	505	30.2	631	10	BF240811
C	30	500.8	29.9	659	14	CF109307
C	31	494	29.5	502	10	AW291800
C	32	489.8	29.3	593	9	AI243720
C	33	485.6	29.0	562	14	N31713
C	34	483.4	28.9	509	12	BM310538
C	35	477.4	28.5	479	12	BM310845
C	36	472.8	28.2	630	14	N31720
C	37	471	28.1	703	12	BP110008
C	38	470	28.1	504	10	AW959892
C	39	465.6	27.8	472	10	AW137803
C	40	463.4	27.7	598	10	BE222114
C	41	448.6	26.8	1398	12	BM912752
C	42	437.6	26.1	593	10	BE376968
C	43	432.4	25.8	434	9	AA406492
C	44	422.6	25.2	429	9	AI066424
C	45	413.2	24.7	973	12	BI154009

ALIGNMENTS

RESULT 1
AK047844
LOCUS AK047844 2450 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:Cl30004M09 product:vascular endothelial growth factor C, full insert sequence.
ACCESSION AK047844
VERSION AK047844.1 GI:26339001
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS

3 Shbata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Obara, E., Watabiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (KISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2450)

JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Atakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kigawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, I., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

JOURNAL
TITLE

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers
1. 2450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C130004M09"
/db_xref="MGI:2413670"
/db_xref="taxon:10090"
/clone="C130004M09"
/issue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="16 days embryo"
175. .1422
/note="unnamed protein product; putative vascular endothelial growth factor C (MGI:109124, GB|NM_009506, evidence: BLASTN, 99%, match=1784)"
/cdon_start=1
/protein_id="BAC33172.1"
/db_xref="GI:26339002"
/translation="MHLCLFLSLACSLIAAALIPSPREAPATVAAFESGLGFSEAPDP

CDS

Query Match 72.1%; Score 1207.6; DB 11; Length 2450;
Best Local Similarity 85.6%; Pred. No. 4.7e-188;
Matches 1418; Conservative 0; Mismatches 219; Indels 19; Gaps 6;
ORIGIN
1 GTCTTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGCTG 60
164 GTCCATCCACCATGCACTTGTCTGTCTCTGTGGCGTGTCTCTGCTCGCGCTG 223
61 CGTGTCTCCCGGCTCCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
224 CGCTGATCCCACTCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 283
121 ACTCTCGGACGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
284 GTTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 343
181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGTAACTCATGACTGTACTACCCAGAT 240
344 AGGAGCAGTTGCGGTCTGTCTCCAGTGTAGTAACTCATGACTGTACTACCCAGAT 403
241 ATTGGAATATGTAAGTGTACGTTAAGGAAAGAGGCTGGCAACATTAACAGAACAG 300
404 ACTGGAATATGTAAGTGTACGTTAAGGAAAGAGGCTGGCAACATTAACAGAACAG 451
301 CCAACTCACTCAAGGAGGAGAGACTATAAATTTGCTGCAGACATTAATATACAG 360
452 CCACCTCAATACCAGGAGGAGAGAGTAAATTTGCTGCAGACATTAATATACAG 511
361 AGATCTTGAATAGTATGATAGTGGAGAGAGAGTAAATTTGCTGCAGACATTAATATACAG 420
512 AGATCTTGAATAGTATGATAGTGGAGAGAGAGTAAATTTGCTGCAGACATTAATATACAG 571
421 GTATAGATGTGGGAGAGAGTGTGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAG 480
572 GTATAGATGTGGGAGAGAGTGTGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAG 631
481 TGTCCGTCTACAGATGTGGGAGTGTCTGCAATAGTGGAGGAGGAGGAGGAGGAGGAGGAG 540
632 TGTCCGTCTACAGATGTGGGAGTGTCTGCAACAGGAGGAGGAGGAGGAGGAGGAGGAG 591
541 GCAAGAGTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
692 GCACAGGTTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 751
601 AACCAAGTAACTCAGTGTTCGCAATACACTTCTCCGATGATGATGATGATGATGATGAT 660
752 AACCAAGTAACTCAGTGTTCGCAATACACTTCTCCGATGATGATGATGATGATGATGAT 811
661 TTTACAGACAGTTCATTTCATTATTATAGACGTTCCCTCCGAGCAACACTACACAGTGT 720
812 TTTACAGACAGTTCATTTCATTATTATAGACGTTCTCTCCGAGCAACACTACACAGTGT 871
721 AGGAGCGGAGCAAGAGCTGCCCGCCCAATTTACATGTGGAATTAATCATATCTGCAGATGCC 780
872 AGGAGCGGAGCAAGAGCTGCCCGCCCAATTTACATGTGGAATTAATCATATCTGCAGATGCC 931
781 TGGCTCAGGAGATTTTATTTCTCTCGGATGCTGGAGATGACTCAACAGATGATGATGCC 840
932 TGGCTCAGGAGATTTTATTTTATTTCAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGA 991
841 ATGACATCTGTGGAGCAAGAGCTGGAGTGAAGAGACCTGTGCAGTGTGTGTGTGCAGAG 900
992 ATGATGTCTGTGGAGCAAGAGCTGGAGTGAAGAGACCTGTGCAGTGTGTGTGTGCAGAG 1051

QY 901 CCGGCTTCCGCTCCAGCTGTGACCCCAAAAGAACTAGACAGAACTCATGCGAGT 960
 Db 1052 GGGGCTTCCGCTTCTAGTTGTGACCCCAAAAGAACTAGACAGCTCATGTCAGT 1111
 QY 961 GTGCTGTAAACAACTCTTCCCGAGCAATGTGGGCCAACCGAGAAATTTGATGAA 1020
 Db 1112 GTGCTGTAAACAACTCTTCCCGAGCAATGTGGGCCAACCGAGAAATTTGATGAA 1171
 QY 1021 ACACATGCCAGTGTGTATGTAAGAACTGCGCCAGAAATCAACCCCTAATCTTGAA 1080
 Db 1172 ATACATGTCAAGTGTATGTAAGAACTGCGCCAGAAATCAACCCCTAATCTTGAA 1231
 QY 1081 AATGTGCTGTCAATGTACAGAACTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
 Db 1232 AATGTGCTGTCAATGTACAGAACTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1291
 QY 1141 ACCACAAACATGAGCTGTATACAGAGCGCCATGACGAACCGCCAGAGCTTTGTGAGC 1200
 Db 1292 ACCACAAACATGAGCTGTATACAGAGCGCCATGACGAACCGCCAGAGCTTTGTGAGC 1351
 QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTGCTCCCTCATATTGGCAAGACCAC 1260
 Db 1352 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTGCTCCCTCATATTGGCAAGACCAC 1411
 QY 1261 AATGAGCTTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTATGGAATACTGTGT 1320
 Db 1412 ATCTCACTAAGATCATACCACTTTTCACTCCACATTTTACTCTCTTGAAGAC--TGT 1469
 QY 1321 TGCCACAGTGAAGCTGTGTGAACAGAGAGACCCCTGTGGGTGCATGCTTAACAAAGACA 1380
 Db 1470 TGCCACATAGCACTGTGTGTGAACAGAGAGCTC--TGTGGGACCATGGTAAACAGAGGCC 1528
 QY 1381 AAGTCTGTCTTCTGGAACCATGGATTAACCTTTACAGAAATGCACTGGAGCTCATCTG 1440
 Db 1529 CAAGTCTGTCTTGTGAACCATGGATTAACCTTTACAGAAATGCACTGGAGCTCATCTG 1586
 QY 1441 CAAAGGCTCTGTGTAAGACTGTTTCTGCCAATGACCAACAGCAAGCAATTTTCTC 1500
 Db 1587 CAAAGGCTCTGTGTAAGACTGTTTCTGCCAATGACCAACAGCAAGCAATTTTCTC 1646
 QY 1501 TTGCTATTTCTTAAAGAACTGCTATATATTTTCTCACTAATAATTTGTTCTGC 1560
 Db 1647 TTGCTATTTCTTAAAGAACTGCTATATATTTTCTCACTAATAATTTGTTCTGC 1705
 QY 1561 ATTCAATTTTATAGCAACAACTATGTAAGAACTGCTATATTTTATATATCAT 1620
 Db 1706 ATTCAATTTTATAGCAACAACTATGTAAGAACTGCTATATTTTATATATCAT 1765
 QY 1621 GCAAAA--TATGTTTAAATAAATAAGAAATTTGTAAT 1655
 Db 1766 GCAAAAATATGTTTAAATAAATAAGAAATTTGTAAT 1801

RESULT 2
 BG677890
 LOCUS 602625584F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4750509 5', linear EST 01-MAY-2001
 DEFINITION mRNA sequence.
 ACCESSION BG677890
 VERSION BG677890.1 GI:13909287
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 998)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL0605 row: g column: 22
 High quality sequence stop: 806.
 Location/Qualifiers
 1..998
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4750509"
 /tissue_type="squamous cell carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Skn4"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

FEATURES source

ORIGIN

Query Match 50.1%; Score 839.4; DB 12; Length 998;
 Best local Similarity 95.4%; Pred. No. 1.1e-127;
 Matches 951; Conservative 0; Mismatches 36; Indels 10; Gaps 8;
 QY 597 CCCAACCCAGTAACATCAGTTTGGCCATACACACTTCTCGCGATGCATGCTAAACTG 656
 Db 1 CCCAACCCAGTAACATCAGTTTGGCCATACACACTTCTCGCGATGCATGCTAAACTG 60
 QY 657 GATGTTTACAGCAACAGTTTCATTATTAGACGTTCCCTGCCAGCAACACTACACAG 716
 Db 61 GATGTTTACAGCAACAGTTTCATTATTAGACGTTCCCTGCCAGCAACACTACACAG 120
 QY 717 TGTACGCGAGCAACAGACCTGCCCCCAATTCATGTAATTAATCACAATCTGCAGA 776
 Db 121 TGTACGCGAGCAACAGACCTGCCCCCAATTCATGTAATTAATCACAATCTGCAGA 180
 QY 777 TGCTCTGCTCAGAGAAATTTATGTTTCTCGGATGCTGGAGATGACATCAACAGATGA 836
 Db 181 TGCTCTGCTCAGAGAAATTTATGTTTCTCGGATGCTGGAGATGACATCAACAGATGA 240
 QY 837 TTCCATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGAGTGTGCTGC 896
 Db 241 TTCCATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGAGTGTGCTGC 300
 QY 897 AGAGCGGGCTTGGGCTCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGTC 956
 Db 301 AGAGCGGGCTTGGGCTCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGTC 360
 QY 957 CAGTGTGTCTGTAAACAACTCTTCCCGAGCAATGTGGGCCCAACCGAGAAATTTGAT 1016
 Db 361 CAGTGTGTCTGTAAACAACTCTTCCCGAGCAATGTGGGCCCAACCGAGAAATTTGAT 420
 QY 1017 GAAACACATGCCAGTGTGTATGTAAGAACTGCTGCCCCAGAAATCAACCCCTAATCT 1076
 Db 421 GAAACACATGCCAGTGTGTATGTAAGAACTGCTGCCCCAGAAATCAACCCCTAATCT 480
 QY 1077 GGAATATGCTGTGAATGTACAGAACTGTCACAGAAATGCTTGTGTTAA--AGGAAGAA 1135
 Db 481 GGAATATGCTGTGAATGTACAGAACTGTCACAGAAATGCTTGTGTTAA--AGGAAGAA 540
 QY 1136 GTTCCACCAACCAATGACGCTGTTTACAGAGCGGCATGTACGAACCCGCCAGAGCTTG 1195
 Db 541 GTTCCACCAACCAATGACGCTGTTTACAGAGCGGCATGTACGAACCCGCCAGAGCTTG 600
 QY 1196 TGAGCCAGATTTTCATATAGTGAAGAGTGTGTTGTCCTTCATATGTCGAAG 1255
 Db 601 TGAGCCAGATTTTCATATAGTGAAGAGTGTGTTGTCCTTCATATGTCGAAG 660
 QY 1256 ACCACAAATGAGCTAAGATTGTACTGTTTCCAGTTTCAT--CGATTTCCTATTATGAAAA 1314
 Db 661 ACCACAAATGAGCTAAGATTGTACTGTTTCCAGTTTCATCCGATTCTCTATTATGAAAA 720

```

QY 1315 CTGTGTTGCCACAGTAGAATGTC-TGTGAACAGAGAGACCCTTGTGGTTCATGCTAAC 1373
Db 721 CTGTGTTGCCACAGTAGAATGTC-TGTGAACAGAGAGAGACCCTTGTGGTTCATGCTAAC 780
QY 1374 AAGACACAAAGTCTGTCTTCCCTGAACC-ATGTGGATACTTTACAGA--AATGGACTGG 1430
Db 781 AAGACACAAAGTCTGTCTTCCCTGAACCAATGGATTTACTTTACAGACATGGACCTGG 840
QY 1431 AGCTACTCTGCAAAAGCCCTCTTGTAAAGACTGGTTTC-TGCCAATGACCAACAGCCA 1489
Db 841 AGCTACTCTGCAAAAGCCCTCTTGTAAAGACCCTTTTCTTGCAATTGACCAACAGCCC 900
QY 1490 AGATTTTCCCTCTGTGATTTCTTTAAAGAGATGACTATATATTTTCCACTTAAAT 1549
Db 901 AGATTTTCCCTCTGTGATTTCTTTAAGA-CATGACTCTCTTAATTTATCCCOCT--TAAT 957
QY 1550 ATTGTTTCTGCATTCATTTTATAGCAACAACTTG 1586
Db 958 TTTGGTTCTGCATTCATTTTATAGCAACAACTTGCTG 994

RESULT 3
BG680919
LOCUS 602628716F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753755 5',
DEFINITION mRNA sequence.
ACCESSION BG680919
VERSION BG680919.1 GI:13912303
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 831)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10613 row: o column: 04
High quality sequence stop: 810.
Location/Qualifiers
1. 831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753755"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-Sport6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
Query Match 46.7%; Score 782.4; DB 12; Length 831;
Best Local Similarity 99.4%; Pred. No. 2.5e-118;
Matches 827; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 43 GTTCTCTGTCGGCGTGCCTGCTCCGGGTCTCGAGAGCGCCCGCGCGCGCG 102
Db 1 GTTCTCTGTCGGCGTGCCTGCTCCGGGTCTCGAGAGCGCCCGCGCGCGCG 60
QY 103 CTTTCGAGTCCGACTCGACTCTCTCGACGCGGAGCCCGAGCGGCGGCGGCTT 162

```

```

Db 61 CTTTCGAGTCCGAGTCTCGACCTCTCGACGCGGAGCCCGAGCGGCGGCGGCTT 120
QY 163 ATGCAAGCAAAAGATCTGGAGGAGCAGTTTACGGTCTGTGTCCAGTGTAGTGAATCTCATGA 222
Db 121 ATGCAAGCAAAAGATCTGGAGGAGCAGTTTACGGTCTGTGTCCAGTGTAGTGAATCTCATGA 180
QY 223 CTGTACTCTACCCAGATATTTGGAAATGTACAAAGTGTACAGTGTACAGTAAAGAAAGAGCGCTGGC 282.
Db 181 CTGTACTCTACCCAGATATTTGGAAATGTACAAAGTGTACAGTGTACAGTAAAGAAAGAGCGCTGGC 240
QY 283 AACATAACAGAGAACAGCCCAACCTCAACTCAAGCAGAGAGAGACTATATAAATTTGCTG 342
Db 241 AACATAACAGAGAACAGCCCAACCTCAACTCAAGCAGAGAGAGACTATATAAATTTGCTG 300
QY 343 CAGCACATTATAATACAGAGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAAT 402
Db 301 CAGCACATTATAATACAGAGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAAT 360
QY 403 GCATGCCACGGAGGTGTGTATAGATGTGGGAGGAGTTTGGAGTCCGACACACACCT 462
Db 361 GCATGCCACGGAGGTGTGTATAGATGTGGGAGGAG- TTGGAGTCCGACACACCT 419
QY 463 TCTTTAAACCTCCATGTGTCTCTACAGATGTGGGGTGTGTCAAATAGTAGAGGGC 522
Db 420 TCTTTAAACCTCCATGTGTCTCTACAGATGTGGGGTGTGTCAAATAGTAGAGGGC 479
QY 523 TGCAGTGCATGAACACAGACAGAGTACCTCAGCAAGAGCTTATTTTGAATTTACAGTGC 582
Db 480 TGCAGTGCATGAACACAGACAGAGTACCTCAGCAAGAGCTTATTTTGAATTTACAGTGC 539
QY 583 CTCTCTCTCAAGGCCCCCAACACAGTAACTCAGTTTTCCTCAATCAGTCTCCCTGCCGAT 642
Db 540 CTCTCTCTCAAGGCCCCCAACACAGTAACTCAGTTTTCCTCAATCAGTCTCCCTGCCGAT 599
QY 643 GCATGTCTAACTGGATGTTTACAGACAGTTCATTCATTTATAGAGTTCCTCCGCGAG 659
Db 600 GCATGTCTAACTGGATGTTTACAGACAGTTCATTCATTTATAGAGTTCCTCCGCGAG 622
QY 703 CAACACTACACAGTGTCCAGCAGCAGAACAGCTCCGCCCAACCAATATCATGTTGGATA 762
Db 660 CAACA-TACACAGTGTCCAGCAGCAGAACAGCTCCGCCCAACCAATATCATGTTGGATA 718
QY 763 ATCATCTCTCAGATGCTGCTCAGGAAGATTTATGTTTTCCTCGGATGCTCGAGATG 822
Db 719 ATCATCTCTCAGATGCTGCTCAGGAAGATTTATGTTTTCCTCGGATGCTCGAGATG 777
QY 823 ACTCAACAGATGGATTCATGACATCTGTGACCAACAAAGAGCTGGATGA 874
Db 778 ACTCAACAGATGGATTCATGACATCTGTGACCAACAAAGAGCTGGATGA 828

RESULT 4
CA488579
LOCUS CA488579
DEFINITION AGENCOURT 10808789 MAPcL Homo sapiens cDNA clone IMAGE:6720351 5',
mRNA sequence.
ACCESSION CA488579
VERSION CA488579.1 GI:24950807
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 858)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```



```

Db      389  TCCAGTTTCATCGATTTCCTATTATGGAATACTGTGTCACAGTAGACCTGTGGA 330
QY      1345  CAGAGAGACCTGTGGGTCCATGCTAAACAAGACAAAGTCTCTTCTCTGTAACCATG 1404
Db      329  CAGAGAGACCTGTGGGTCCATGCTAAACAAGACAAAGTCTCTTCTGTAACCATG 270
QY      1405  TGGTAACCTTTACAGAAATGACCTGGAGCTCATCTGCAAAAGGCTCTTGTAAAGACTGG 1464
Db      269  TGGTAACCTTTACAGAAATGACCTGGAGCTCATCTGCAAAAGGCTCTTGTAAAGACTGG 210
QY      1465  TTTTCTGCCAATGACCAACAGCCCAAGATTTCCTCTTGTGATTTCTTTAAAGAAATGAC 1524
Db      209  TTTTCTGCCAATGACCAACAGCCCAAGATTTCCTCTTGTGATTTCTTTAAAGAAATGAC 150
QY      1525  TATATAATTTATTTCCACTAAATAATTTGTTCTTGCATTTATTTTATAGCAACAAT 1584
Db      149  TATATAATTTATTTCCACTAAATAATTTGTTCTTGCATTTATTTTATAGCAACAAT 90
QY      1585  TGGTAAAACTCAGCTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATATAATG 1644
Db      89  TGGTAAAACTCAGCTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATATAATG 30
QY      1645  AAAATTGATTTATAAAAAAATAAAAAA 1673
Db      29  AAAATTGATTTATAAAAAAATAAAAAA 1

```

RESULT 7

AW195617/c

LOCUS

DEFINITION

xh84g11.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone

IMAGE:2701220 3', similar to SW:VEGC HUMAN P49767 VASCULAR

ENDOTHELIAL GROWTH FACTOR C PRECURSOR ;, mRNA sequence.

AW195617

AW195617.1 GI:8474775

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 808)

NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 498.

Location/Qualifiers

FEATURES

source

1..808

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="IMAGE:2701220"

/lab_host="DH10B"

/clone_lib="Soares NFL T GBC S1"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site_1: Not I; Site_2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBHL19W, testis NHT, and B-cell

NCI CGAP GCBI) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

1.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Fatima Bonaldo."

ORGIN

Query Match 41.9%; Score 702; DB 10; Length 808;
 Best Local Similarity 94.4%; Pred. No. 3.7e-105;
 Matches 760; Conservative 0; Mismatches 41; Indels 4; Gaps 3;

```

QY      855  CCAACAAGAGCTGGATGAAGAGACCTGTAGTGTGTGTCAGAGCGGGCTTGGCGCT 914
Db      808  CCAACAAGAGCTGGAGAGACAGACCGGTGAGTGTGTGAAGACAGGGGTTTCGGCTG 749
QY      915  GCCAGCTGTGGACCCCAACAAGAACTAGACAGAACTCATGCCAGTGTCTGTGAAAAAC 974
Db      748  CCAGCTGTGGCCCCCACCACCAAGAACTAGACAGTACTCAATGCCAGTGTCTGTGAAAAAC 689
QY      975  AAACCTTTCCCGAG-CCAATGTGGGCCCAACC--GAGAAATTTGATGAAACACATGCCAG 1031
Db      688  CAACTCTTCCCGAGCCCAATGTGGGCCCAACCAGGAATTTTGTGAAACACATGCCAG 629
QY      1032  -TGTTGATGTAAGAAACCTGCCCCAGAAATCAACCCCTTAAATCTCGGAAAAATGTGCCCTG 1090
Db      628  TTGTGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTTAAATCTGGAATGTGCCCTG 569
QY      1091  TGAATGTACAGAAAGTCCACAGAAATGCTTGTGTAAGAAAGAAAGAGTCCACCAACAAC 1150
Db      568  TGAATGTACAGAAAGTCCACAGAAATGCTTGTGTAAGAAAGAAAGAGTCCACCAACAAC 509
QY      1151  ATGCAGCTGTACAGACGCCCATGTACGAACCGCCAGAGGCTTGTGAGCAGGATTTTC 1210
Db      508  ATGCAGCTGTACAGACGCCCATGTACGAACCGCCAGAGGCTTGTGAGCAGGATTTTC 449
QY      1211  ATATAGTGAAGAGTGTGTGTTGTGTCCTTCAATATTTGGCAAGACCAACAATGAGCTA 1270
Db      448  ATATAGTGAAGAGTGTGTGTTGTGTCCTTCAATATTTGGCAAGACCAACAATGAGCTA 389
QY      1271  AGATTGTACTGTTTCCAGTTTCATCGATTTCTATTATGGAAGAACTGTGTTGCCACAGTA 1330
Db      388  AGATTGTACTGTTTCCAGTTTCATCGATTTCTATTATGGAAGAACTGTGTTGCCACAGTA 329
QY      1331  GAACTGTCTGTGAACAGAGAGACCCCTTGTGGTCCATGCTAACCAAGACCAAAAGTCTGTC 1390
Db      328  GAACTGTCTGTGAACAGAGAGACCCCTTGTGGTCCATGCTAACCAAGACCAAAAGTCTGTC 269
QY      1391  TTTCTGAAACCAATGGAATACTTTACAGAAATGAGCTGGAGCTCATCTGCAAAAGGCT 1450
Db      268  TTTCTGAAACCAATGGAATACTTTACAGAAATGAGCTGGAGCTCATCTGCAAAAGGCT 209
QY      1451  CTTGTAAAGACTGTTTCTGCAATGACCAACAGCAAGATTTTCTCTTGTGATTTTC 1510
Db      208  CTTGTAAAGACTGTTTCTGCAATGACCAACAGCAAGATTTTCTCTTGTGATTTTC 149
QY      1511  TTTAAAGAAGTACATATAATTTATTTCCACTAAAAATATTGTTTCTGCAATTTCAATTTT 1570
Db      148  TTTAAAGAAGTACATATAATTTATTTCCACTAAAAATATTGTTTCTGCAATTTCAATTTT 89
QY      1571  ATAGCAACAACAATTTGGTAAACCTCACTGTGATCAATATTTTATATCATGCAAAATATG 1630
Db      88  ATAGCAACAACAATTTGGTAAACCTCACTGTGATCAATATTTTATATCATGCAAAATATG 29
QY      1631  TTTAAATAAATAAATGAAAAATTTGTAAT 1655
Db      28  TTTAAATAAATAAATGAAAAATTTGTAAT 4

```

RESULT 8

AA923193/c

LOCUS

DEFINITION

om48d01.s1 Soares NFL T GBC S1 Homo sapiens cDNA clone

IMAGE:1544257 3', similar to SW:VEGC HUMAN P49767 VASCULAR

ENDOTHELIAL GROWTH FACTOR C PRECURSOR ;, mRNA sequence.

AA923193

AA923193.1 GI:3070502

EST.

Homo sapiens (human)

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 717)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 785 Std Error: 0.00
Seq primer: -40ml3 fwd. EF from Amersham
High quality sequence stop: 476.
Location/Qualifiers
1..717
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1544257"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: pooled; Vector: pVT3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

REFERENCE
1 (bases 1 to 717)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2601 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 457.
Location/Qualifiers
1..728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2567826"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: pooled; Vector: pVT3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 41.7%; Score 698.6; DB 9; Length 717;
Best Local Similarity 99.6%; Pred. No. 1.4e-104;
Matches 704; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
948 AACCTCATGCTGTGCTGTAAACCAAACTCTTCCACGCAATGTGGGGCCACCGA 1007
714 AATCATGCTGTGCTGTGTAAACCAAACTCTTCCACGCAATGTGGGGCCACCGA 655
1008 GAATTTGATGAACACATCCAGTGTGTATGTAAGAACCTGCCCGCAATCAACCC 1067
654 GAATTTGATGAACACATCCAGTGTGTATGTAAGAACCTGCCCGCAATCAACCC 595
1068 CTAAATCTCGAAATGTGCTGTGAATGTACAGAAATCCACAGAAATGCTTTTAAAA 1127
594 CTAAATCTCGAAATGTGCTGTGAATGTACAGAAATCCACAGAAATGCTTTTAAAA 535
1128 GGAAGAAGATTCCACCAACCAACATGCTGTGTTACAGAGCGCATGTACGACCCCGCAG 1187
534 GGAAGAAGATTCCACCAACCAACATGCTGTGTTACAGAGCGCATGTACGACCCCGCAG 475
1188 AAGCTTTGTAGCCAGAGATTTCATATAGTGAAGAGTGTGTGTGTGTCCTTCATAT 1247
474 AATGCTTTGTAGCCAGGATTCATATAGTGAAGAGTGTGTGTGTGTCCTTCATAT 415
1248 TGGCAAGACCCACAAATGAGTGAATGTGTTTCCAGTTTCATGATTTTCATTATTA 1307
414 TGGAAAGACCAACAAATGAGTGAATGTGTTTCCAGTTTCATGATTTTCATTATTA 355
1308 TGGAAATCTGTGTGTCACAGTAGAATGCTGTGTGAACAGAGAGACCTTGTGGGTCCAT 1367
354 TGGAAATCTGTGTGTCACAGTAGAATGCTGTGTGAACAGAGAGACCTTGTGGGTCCAT 295
1368 GCTAACAAAGACAAAGTCTGCTTCTTCCAGCATGTGTGATACCTTTACAGAAATGGAC 1427
294 GCTAACAAAGACAAAGTCTGCTTCTTCCAGCATGTGTGATACCTTTACAGAAATGGAC 235
1428 TGGAGCTCATGTGCAAAAGGCTCTTGTGAAGACTGGTTTTCTGCCAATGACCAACAGC 1487

Db 234 TGGAGCTCATGTGCAAAAGGCTCTTGTGAAGACTGGTTTTCTGCCAATGACCAACAGC 175
Qy 1488 CAAGATTTCCTCTGTGTGATTCTTTAAAGAAAGTACTATATAATTTATTTCCACTAAA 1547
Db 174 CAAGATTTCCTCTGTGTGATTCTTTAAAGAAAGTACTATATAATTTATTTCCACTAAA 115
Qy 1548 ATATGTTTCGCAATTCATTTTATAGCAACAAATGGTAAACCTCCTGTGATCAAT 1607
Db 114 ATATGTTTCGCAATTCATTTTATAGCAACAAATGGTAAACCTCCTGTGATCAAT 55
Qy 1608 ATTTTATATCATGCAAAATATGTTTAAATAAATAAATGAAATTTGATTTATAAA 1661
Db 54 ATTTTATATCATGCAAAATATGTTTAAATAAATAAATGAAATTTGATTTATAAA 1

RESULT 9
AW070215/c
LOCUS
DEFINITION
xa09c10.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2567826 3', similar to SW:VEGC_HUMAN P49767 VASCULAR
ENDOTHELIAL GROWTH FACTOR C PRECURSOR 1, mRNA sequence.
AW070215
AW070215.1 Gi:6025213
EST.
ACCESSION
AW070215
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 728)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2601 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 457.
Location/Qualifiers
1..728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2567826"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: pooled; Vector: pVT3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 39.9%; Score 667.2; DB 9; Length 728;
Best Local Similarity 97.7%; Pred. No. 1.9e-99;
Matches 708; Conservative 0; Mismatches 14; Indels 3; Gaps 3;
Qy 933 AAAGAACTAGACAGAACTCATGCCAGTGTCTGTGTAATAAACAACCTTCCCGAGCAA 992
Db 728 AAAGAACTAGACAGAACTCATGCCAGTGTCTGTGTAATAAACAACCTTCCCGAGCAA 670
Qy 993 TGTGGGGCCAAACCGAGAAATTTGATGAAACACATGCCAG-TGTTGATGTAAGAAACCTG 1051

669	GGTGGGCAACCCGAGATTTGNATGAAAAACACATGCCAGNTGTGTATGTAAAAAGAACCTG	610
1052	CCCAGAAATCAACCCCTAAATCCTGGAAAAATGTGCCTGTGAATGTACAGAAAGTCCACA	1111
609	CCCAGAAATCAACCCCTAAATCCTGGAAAAATGTGCCTGTGAATGTACAGAAAGTCCACA	550
1112	GAATGCTGTGTAAAGGGAAGAGTTCCACCCACCAACATGCAGCTGTTTACAGACGGCC	1171
549	GAATGCTGTGTAAAGGGAAGAGTTCCACCCACCAACATGCAGCTGTTTACAGACGGCC	490
1172	ATGTAGCAACCGCCAGAGGCTTTGTGAGCCAGGATTTTTCATATAGTGAAGAAGTGTGTG	1231
489	ATGTAGCAACCGCCAGAGGCTTTGTGAGCCAGGATTTTTCATATAGTGAAGAAGTGTGTG	430
1232	TTGTGTCCTTCATATTGGCAAGACACAAAATGAGCTAAGATTGTACTGTTTCCAGTT	1291
429	TTGTGTCCTTCATATTGGCAAGACACAAAATGAGCTAAGATTGTACTGTTTCCAGTT	370
1292	CATCGATTTTCTATTATGGAATACTGTGTTGCCACAGTAGAATGCTGTGGAACAGAGAG	1351
369	CATCGATTTTCTATTATGGAATACTGTGTTGCCACAGTAGAATGCTGTGGAACAGAGAG	310
1352	ACCTTGTGGTCCATGCTTAACAAAGACAAAAGTCTGTCTTTCCTGAACCATGTGGATAA	1411
309	ACCTTGTGGTCCATGCTTAACAAAGACAAAAGTCTGTCTTTCCTGAACCATGTGGATAA	250
1412	CTTTACAGAAATGGACTGG-AGCTCATCTGCAAAAGGCCTCTTGTAAAGACTGGTTTCT	1470
249	CTTTACAGAAATGGACTGGAAAGTCACTGCAAAAGGCCTCTTGTAAAGACTGGTTTCT	150
1471	GCCAAATGACCAACAGCCAGATTTTCTCTGTGATTTCTTTAAAGAATGACTATATA	1530
189	GCCAAATGACCAACAGCCAGATTTTCTCTGTGATTTCTTTAAAGAATGACTATATA	130
1531	ATTATTTCACCTAAAAATATTGTTTCTGCAATCTATTTTATAGCAACAAATTTGGTAA	1590
129	ATTATTTCACCTAAAAATATTGTTTCTGCAATCTATTTTATAGCAACAAATTTGGTAA	70
1591	AACCTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAAAATGAAAAATT	1650
69	AACCTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAAAATGAAAAATT	10
1651	GTATT 1655	
9	GTATT 5	

RESULT 10	
LOCUS	BQ005831/c
DEFINITION	UI-H-EDO-avr-n-11-0-UI.sl NCI CGAP_EDO Homo sapiens CDNA clone IMAGE:5842762 3', mRNA sequence.
ACCESSION	BQ005831
VERSION	BQ005831.1
KEYWORDS	GI:19730731
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 662)
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapsb@mail.nih.gov Tissue Procurement: Dr. Jose Mercuende CDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: http://image.lnl.gov The following repetitive elements were found in this cDNA

```

sequence: 1-53, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             Location/Qualifiers
     1..662
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:5842762"
         /tissue_type="Chondrosarcoma"
         /dev_stage="Adult"
         /lab_host="DH10B (Life Technologies)"
         /clone_lib="NCI CGAP ED0"
         /note="Organ: Left Pubic Bone; Vector: pT7T3-Pac
         (Pharmacia) with a modified polylinker; Site 1: EcoR I;
         Site 2: Not I; NCI CGAP_ED0 is a cDNA library containing
         the following tissue(s): Chondrosarcoma cell line C59. The
         library was constructed according to Bonaldo, Lennon and
         Soares, Genome Research, 6:791-806, 1996. First strand
         cDNA synthesis was primed with an oligo-dT primer
         containing a Not I site. Double stranded cDNA was ligated
         to an EcoR I adaptor, digested with Not I, and cloned
         directionally into pT7T3-Pac vector. The oligonucleotide
         used to prime the synthesis of first-strand cDNA contains
         a library tag sequence that is located between the Not I
         site and the (dT)18 tail. The sequence tag for this
         library is GCTCAAGGCT.
         TAG_TISSUE=chondrosarcoma
         TAG_LIB=UI-H-ED0
         TAG_SEQ=CGTCAAGGCT

```

ORIGIN	Query Match	39.2%;	Score 656.2;	DB 12;	Length 662;
	Best Local Similarity	99.5%;	Pred. No. 1.2e-97;		
	Matches 658;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1014	GATGAAACACATGCCAGTGTGTATCTAAAGAACTGCCCCAGAAATCAACCCCTAAAT	1073		
Db	662	GATGAAACACATGCCAGTGTGTATCTAAAGAACTGCCCCAGAAATCAACCCCTAAAT	603		
QY	1074	CTGTGAAATATGTGCTGTGAAATGTATACAGAAAGTCCACAGAAATGCTTTAAAGGAAAG	1133		
Db	602	CTGTGAAATATGTGCTGTGAAATGTATACAGAAAGTCCACAGAAATGCTTTAAAGGAAAG	543		
QY	1134	AAAGTCCACACCAACATGTCAGCTGTTACAGACGCCCATGTACGAAACCGCCAGAAAGGCT	1193		
Db	542	AAAGTCCACACCAACATGTCAGCTGTTACAGACGCCCATGTACGAAACCGCCAGAAAGGCT	483		
QY	1194	TGTGAGCCAGGATTTTCATATAGTGNAGAAAGTGTGTCGTTGTGTCCTTCATATTTGGCAA	1253		
Db	482	TGTGAGCCAGGATTTTCATATAGTGAAGAAAGTGTGTCGTTGTGTCCTTCATATTTGGAAA	423		
QY	1254	AGACCACAAATGAGCTAAGATGTACTGTTTTCCAGTTCATCGAATTTCTATTATGAAA	1313		
Db	422	AGACCACAAATGAGCTAAGATGTACTGTTTTCCAGTTCATCGAATTTCTATTATGAAA	363		
QY	1314	ACTGTGTTGCCACAGTGAACCTCTGTGAAACAGAGAGACCTGTGGGTGCCATGCTAAC	1373		
Db	362	ACTGTGTTGCCACAGTGAACCTCTGTGAAACAGAGAGACCTTTGGGTGCCATGCTAAC	303		
QY	1374	AAAGACAAAGTCTGTCTTCTCGAACCACTGTGGATACTTTACAGAAATGGACTGGAGC	1433		
Db	302	AAAGACAAAGTCTGTCTTCTCGAACCACTGTGGATACTTTACAGAAATGGACTGGAGC	243		
QY	1434	TCACTGCAAAAGGCTCTTTGTAAAGACTGGTTTTCTGCCAATGACCAACACGCCAAGAT	1493		
Db	242	TCACTGCAAAAGGCTCTTTGTAAAGACTGGTTTTCTGCCAATGACCAACACGCCAAGAT	183		
QY	1494	TTTCTCTTTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTG	1553		
Db	182	TTTCTCTTTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTG	123		
QY	1554	TTTCTGCATCTCATTTTTTATAGCAACAACTAATTTGGTAAAACTCACGTGTGATCAATATTTT	1613		

Db 122 TTTCGCAATCAATTTATAGCAACAACAATGGTAAACCTCACTGTGATCAATATTTT 63
 1614 ATATCATGCAAAATATGTTTAAATAAATAAATGAAATTCGTTATTAATAAATAAATAAATAA 1673
 Db 62 ATATCATGCAAAATATGTTTAAATAAATAAATGAAATTCGTTATTAATAAATAAATAAATAA 3
 Qy 1674 A 1674
 Db 2 A 2

RESULT 11
 LOCUS BU619991/c
 DEFINITION UI-H-FL1-bfv-p-05-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
 ACCESSION UI-H-FL1-bfv-p-05-0-UI 3', mRNA sequence.
 VERSION BU619991
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 657)
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 Cloning Distribution: Cloning Distribution Information can be obtained
 from Dr. M. Bento Soares, bento-soares@iowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 1-48, >AT rich/Low complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes

FEATURES

location/Qualifiers
 1..657
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FL1-bfv-p-05-0-UI"
 /tissue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP FL1"
 /note="Organ: Chondrosarcoma; Vector: pTTT3-Pac
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
 derived from a pool of mRNA obtained from 4 cell lines
 from grade III chondrosarcoma tissues. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pTTT3-Pac vector. The oligonucleotide used to prime
 the synthesis that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 GAGGTCGGTG. The cell lines were provided by Dr. James
 Martin from the University of Iowa.
 TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG_LIB=UI-H-FL1
 TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 38.8%; Score 649; DB 13; Length 657;

Best Local Similarity 99.2%; Pred. No. 1.8e-96;
 Matches 652; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1014 GATGAAAACACATGCCAGTGTGTATGTAAGAAAGAACTGCCAGAAATCAACCCCTAAAT 1073
 Db 657 GATGAAAACACATGCCAGTGTGTATGTAAGAAAGAACTGCCAGAAATCAACCCCTAAAT 598
 Qy 1074 CTTGGAATAATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAG 1133
 Db 597 CTTGGAATAATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAG 538
 Qy 1134 AGTTCCACCAACATGCAGCTGTGTACAGAGCGCCATGTACGACCGCCAGAGGCT 1193
 Db 537 AGTTCCACCAACATGCAGCTGTGTACAGAGCGCCATGTACGACCGCCAGAGGCT 478
 Qy 1194 TGTGAGCCAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCCTTCATATTCGCAA 1253
 Db 477 TGTGAGCCAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCCTTCATATTCGCAA 418
 Qy 1254 AGACCAAAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTCTATATGGAAA 1313
 Db 417 AGACCAAAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTCTATATGGAAA 358
 Qy 1314 ACTGTGTTGCCACAGTAGAACTGTCTGCAACAGAGAGACCTTGTGGGTCCATGCTAAC 1373
 Db 357 ACTGTGTTGCCACAGTAGAACTGTCTGCAACAGAGAGACCTTGTGGGTCCATGCTAAC 298
 Qy 1374 AAAGCAAAAGTCTGTTTCTCTGCAACCAATGCGATACTTTACAGAAATGAGCTGGAGC 1433
 Db 297 AAAGCAAAAGTCTGTTTCTCTGCAACCAATGCGATACTTTACAGAAATGAGCTGGAGC 238
 Qy 1434 TCATCTGCAAAAGGCTCTTGTAAAGACTGTTTCTGCAATGACCAACAGCCCAAGAT 1493
 Db 237 TCATCTGCAAAAGGCTCTTGTAAAGACTGTTTCTGCAATGACCAACAGCCCAAGAT 178
 Qy 1494 TTTCTCTTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTG 1553
 Db 177 TTTCTCTTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTG 118
 Qy 1554 TTTCTGCAATTTTATAGCAACCAATGTTGTTAAACTCAGTGTGATCAATATTTT 1613
 Db 117 TTTCTGCAATTTTATAGCAACCAATGTTGTTAAACTCAGTGTGATCAATATTTT 58
 1614 ATATCATGCAAAATATGTTTAAATAAATAAATGAAATTCGTTATTAATAAATAAATAA 1670
 57 ATATCATGCAAAATATGTTTAAATAAATAAATGAAATTCGTTATTAATAAATAAATAA 1

RESULT 12

BE218812/c
 LOCUS BE218812
 DEFINITION BE218812 637 bp mRNA linear EST 03-JUL-2000
 hv45a10.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3176346 3,
 similar to SW.VEGC HUMAN P49767 VASCULAR ENDOTHELIAL GROWTH FACTOR
 C PRECURSOR ;, mRNA sequence.
 BE218812
 VERSION BE218812.1 GI:8906130
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 637)
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov

Seq primer: -40UP from Gibco
High quality sequence stop: 449.

FEATURES

Location/Qualifiers
1..637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3176346"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pTTT3D-Pac (Pharmacia) with a modified polylinker; plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 141920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 36.8%; Score 616; DB 10; Length 637;
Best Local Similarity 99.1%; Pred. No. 4.7e-91;
Matches 630; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1020 AACACATGCCAGTGTGTATGTAAGAGAACCTGCCCGAGAAATCAACCCCTAAATCCTCGA 1079
Db 637 AACACATGCCAGTGTGTATGT-AAAGAACCTGCCCGAGAAATCAACCCCTAAATCCTCGA 579
QY 1080 AATGTGCTGTGAATGTACAGAAATCCACAGAAATGCTGTAAAGAGGAAGAGTTC 1139
Db 578 AATGTGCTGTGAATGTACAGAAATCCACAGAAATGCTGTAAAGAGGAAGAGTTC 519
QY 1140 CACCACCAACATGAGCTGTACAGACGGCCATGTACGACCGCAGAGCTTGTGAG 1199
Db 518 CACCACCAACATGAGCTGTACAGACGGCCATGTACGACCGCAGAGCTTGTGAG 459
QY 1200 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTCCTTCATATATGGCAAGACCA 1259
Db 458 CCAGGATTTTCATATAGTGAAGAGTGTGTGTGTCCTTCATATATGGCAAGACCA 399
QY 1260 CAAATGAGCTAGATGTACTGTTTCCAGTTCATCGATTTTCTATTATGGAAGTCTG 1319
Db 398 CAAATGAGCTAGATGTACTGTTTCCAGTTCATCGATTTTCTATTATGGAAGTCTG 339
QY 1320 TTGCCACAGTAGAATCTGTGTGAACAGAGAGACCCCTGTGGTCCATGCTAACCAAGAC 1379
Db 338 TTGCCACAGTAGAATCTGTGTGAACAGAGAGACCCCTGTGGTCCATGCTAACCAAGAC 279
QY 1380 AAAAGTCTCTTCTCTGACCATGTGGTATACCTTTACAGAAATGAGCTGGAGCTCATCT 1439
Db 278 AAAAGTCTCTTCTCTGACCATGTGGTATACCTTTACAGAAATGAGCTGGAGCTCATCT 219
QY 1440 GCAAAAGGCTCTTGTAAAGAGCTGTTTCTGCCATGACCAAGCAGCAAGTTCCT 1499
Db 218 GCAAAAGGCTCTTGTAAAGAGCTGTTTCTGCCATGACCAAGCAGCAAGTTCCT 159
QY 1500 CTTGTGATTTCTTAAAGAGCTATATATATTTTCCACTAAATAATGTTCCTG 1559
Db 158 CTTGTGATTTCTTAAAGAGCTATATATATTTTCCACTAAATAATGTTCCTG 99
QY 1560 CATTCAATTTTATAGCAACCAAGTGGTAAACCTCAGCTGTGATCAATATTTTATATCA 1619
Db 98 CATTCAATTTTATAGCAACCAAGTGGTAAACCTCAGCTGTGATCAATATTTTATATCA 39
QY 1620 TGCATAATATGTTTAAATAATGAAATTTGATT 1655
Db 38 TGCATAATATGTTTAAATAATGAAATTTGATT 3

RESULT 13

CB854820/c
LOCUS
DEFINITION
UI-CF-EC1-ack-o-03-0-UI-s1 UI-CF-EC1 Homo sapiens cDNA clone
UI-CF-EC1-ack-o-03-0-UI 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB854820 622 bp mRNA linear EST 22-APR-2003
CB854820.1 GI:30045199
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 622)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB

University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171

Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA
sequence: 1-53, >AT-richLow_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers
1..622
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-ack-o-03-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/notes="Organ: Lung; Vector: pTTT3-Pac (Pharmacia) with a modified polylinker; Site_1: Ecor I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pTTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGCTTAC.
TAG_SEQ=None found"

Query Match 36.7%; Score 614.6; DB 14; Length 622;
Best Local Similarity 99.4%; Pred. No. 7.9e-91;
Matches 617; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1054 CCAGAAATCAACCCCTAAATCCTGGAAATGTCCCTGTGTAATGTACAGAAAGTCCACAGA 1113
Db 622 CCAGAAATCAACCCCTAAATCCTGGAAATGTCCCTGTGTAATGTACAGAAAGTCCACAGA 563

```

QY 1114 AATGCTTGTAAAGAAAGAGTTCCACACCAACATGAGCTGTTTACAGAGCGGCAT 1173
Db 562 AATGCTTGTAAAGAAAGAGTTCCACACCAACATGAGCTGTTTACAGAGCGGCAT 503
QY 1174 GTACGAACCGCCAGAGGCTTGTGAGCAGAGATTTTCATATAGTGAAGAGTGTGCTT 1233
Db 502 GTACGAACCGCCAGAGGCTTGTGAGCAGAGATTTTCATATAGTGAAGAGTGTGCTT 443
QY 1234 GTGTCCCTTCATATTGGCAAGACACACAAATGAGCTAAGATTGTACTGTTTCCAGTTCA 1293
Db 442 GTGTCCCTTCATATTGGCAAGACACACAAATGAGCTAAGATTGTACTGTTTCCAGTTCA 383
QY 1294 TCGATTTTCTATTATGGAAGAACTGTGTCACACAGTAACTGTCTGTGAACAGAGAGAC 1353
Db 382 TCGATTTTCTATTATGGAAGAACTGTGTCACACAGTAACTGTCTGTGAACAGAGAGAC 323
QY 1354 CTTTGGGTCCATCTCAACAAAGACAAAGTCTGTTTCTTCTGAAACCAATGGAATACT 1413
Db 322 CTTTGGGTCCATCTCAACAAAGACAAAGTCTGTTTCTTCTGAAACCAATGGAATACT 263
QY 1414 TTACAGAAATGAGCTGAGCTCATCTGCAAGAGGCTCTGTAAGACTGTTTCTGCCC 1473
Db 262 TTACAGAAATGAGCTGAGCTCATCTGCAAGAGGCTCTGTAAGACTGTTTCTGCCC 203
QY 1474 AATGACCAACAGCCAGAGATTTCTCTTGTGANTTTCTTTAAAGAAATGACTATATAAT 1533
Db 202 AATGACCAACAGCCAGAGATTTCTCTTGTGANTTTCTTTAAAGAAATGACTATATAAT 143
QY 1534 TATTTCCACTAAATATGTTTCTGCAATCATTTTATAGCAACACAAATGGAATAAC 1593
Db 142 TATTTCCACTAAATATGTTTCTGCAATCATTTTATAGCAACACAAATGGAATAAC 83
QY 1594 TCACGTGATCAATATTTTATCATGCAAAATATGTTTAAATAAAATGAAATTTGTA 1653
Db 82 TCACGTGATCAATATTTTATCATGCAAAATATGTTTAAATAAAATGAAATTTGTA 23
QY 1654 TTTATAAAAAAATAAAAAA 1674
Db 22 TTGTAATAAAAAAATAAAAAA 2

RESULT 14
BUS10633
LOCUS
DEFINITION
AGNCOURT 10118034 NIH MGC 134 Mus musculus cDNA clone
IMAGE:6505339 5', mRNA sequence.
ACCESSION
BUS10633
VERSION
BUS10633.1 GI:22816866
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 916)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M14066 row: a column: 20
High quality sequence stop: 620.
Location/Qualifiers
1..916
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
FEATURES
source

```

```

/clon="IMAGE:6505339"
/issue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clon_lib="NIH MGC 134"
/notes="Vector: pCMV-Sport6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by Resgen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 36.0%; Score 603; DB 13; Length 916;
Best Local Similarity 84.5%; Pred. No. 5.7e-89;
Matches 731; Conservative 0; Mismatches 115; Indels 19; Gaps 4;

QY 116 ACTCGACCTCTCGAGCGCGAGCCGACGCGGGGAGGCGCCAGGCTTATGCAAGCAAGA 175
Db 1 ACTGGGCTTCTCGGAAGCGGAGCCGACGCGGGGCGAGGTCAAGGCTTTTGAAGGCAAGA 60
QY 176 TCTGGAGGAGCAGTTACGGTCTGTGTCAGTGTAGTGAACCTCATGCTACTCTACCC 235
Db 61 CTTGGAGGAGCAGTTTCCGGTCTGTGTCAGCGTAGATGATGTCTGTCTGTACCC 120
QY 236 AGAATATTGGAATAATGTACAAGTGTCAAGTAAAGAAAGGAGGCTGGCAACATAACAGAGA 295
Db 121 AGACTACTGGAATAATGTACAAGTGTCAAGTGTCAAGTAAAGGCGGCTGGC-----A 168
QY 296 ACAGGCAACCTCACTCAAGACAGAGAGACTATATAATTTGCTGAGCAGACTATATAA 355
Db 169 GCAGGCCACCTCACTCAAGTGTCAAGGAGGAGGAGTGTAAATTTGCTGTGCACTATATAA 228
QY 356 TACAGAGATCTTTGAAAGTATTGATAATGAGTGAAGAAAGACTCAATGATGCCACGGGA 415
Db 229 CACAGAGATCTTGAAGTATTGATAATGAGTGAAGAAAGACTCAATGATGCCACGGGA 288
QY 416 GGTGTGTATAGATGTGGGAGAGAGTTTGGAGTCCGCAACAAACACTTCTTTAAACCTCC 475
Db 289 GGTGTGTATAGATGTGGGAGAGAGTTTGGAGTCCGCAACAAACACTTCTTTAAACCTCC 348
QY 476 ATGTGTGTCGCTACAGATGTGGGAGTGTGTCGAATAGTGAAGGCTGCACTGATGAA 535
Db 349 ATGTGTGTCGCTACAGATGTGGGAGTGTGTCGAATAGTGAAGGCTGCACTGATGAA 408
QY 536 CACAGCAGAGCTACCTCAGCAAGACGTTATTGAAATACAGTGCCTCTCTCTCAAG 595
Db 409 CACAGCAGAGTACCTCAGCAAGACGTTATTGAAATACAGTGCCTCTCTCTCAAG 468
QY 596 CCCCACACAGTAAACATCAGTTTGGCAATCAGACTTCTGCGCATGCTGCTAAACT 555
Db 469 CCCCACACAGTAAACATCAGTTTGGCAATCAGACTTCTGCGCATGCTGCTAAACT 528
QY 656 GGATGTTTACAGACAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACACTACCACA 715
Db 529 GGATGTTTACAGACAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACACTACCACA 588
QY 716 GTGTGAGGAGCGAAGACAGCTGCCACCACTTACATGTGGAATATCATCATCTGAG 775
Db 589 GTGTGAGGAGCGAAGACAGCTGCCACCACTTACATGTGGAATATCATCATCTGAG 648
QY 776 ATGCTGCTGAGTCAAGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGG 835
Db 649 ATGCTGCTGAGTCAAGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGG 708
QY 836 ATTCCATGACATCTGTGAGCAACAAACAGAGCTGGATGAAGAGACC--TGTGAGTGTCT 893
Db 709 ATTCCATGATGCTGTGAGCAACCAACAGAGCTGGATGAAGAGACACCTGTCTCATGTGTC 768
QY 894 TGCAGAGGCGGCTTCCGCTGCCAGCTGTGG-----ACCCACAAAGAACTAGACA--CAAG 948
Db 769 TGCAGAGGCGGCTTCCGCTGCCAGCTGTGG-----ACCCACAAAGAACTAGATTAGAG 828
QY 949 ACTCATGCCAGTGTCTGTGTAATAA 973
Db 829 ACTCATGCCAGGTTGGCTGTGTAATAA 853

```

